

Comparison of the prediction of Indian monsoon low-pressure systems by Subseasonal-to-Seasonal prediction models

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open access

Deoras, A., Hunt, K. M. R. ORCID: <https://orcid.org/0000-0003-1480-3755> and Turner, A. G. ORCID: <https://orcid.org/0000-0002-0642-6876> (2021) Comparison of the prediction of Indian monsoon low-pressure systems by Subseasonal-to-Seasonal prediction models. *Weather and Forecasting*, 36 (3). pp. 859-877. ISSN 0882-8156 doi: <https://doi.org/10.1175/WAF-D-20-0081.1> Available at <https://centaur.reading.ac.uk/96307/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1175/WAF-D-20-0081.1>

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Comparison of the Prediction of Indian Monsoon Low Pressure Systems by Subseasonal-to-Seasonal Prediction Models

AKSHAY DEORAS,^a KIERAN M. R. HUNT,^{a,b} ANDREW G. TURNER,^{b,a}

^a *Department of Meteorology, University of Reading, Reading, United Kingdom*

^b *National Centre for Atmospheric Science, University of Reading, Reading, United Kingdom*

(Manuscript received 22 May 2020, in final form 5 February 2021)

ABSTRACT: This study analyzes the prediction of Indian monsoon low pressure systems (LPSs) on an extended time scale of 15 days by models of the Subseasonal-to-Seasonal (S2S) prediction project. Using a feature-tracking algorithm, LPSs are identified in 11 S2S models during a common reforecast period of June–September 1999–2010, and then compared with 290 and 281 LPSs tracked in ERA-Interim and MERRA-2 reanalysis datasets. The results show that all S2S models underestimate the frequency of LPSs. They are able to represent transits, genesis, and lysis of LPSs; however, large biases are observed in the Australian Bureau of Meteorology, China Meteorological Administration (CMA), and Hydrometeorological Centre of Russia (HMCR) models. The CMA model exhibits large LPS track position error and the intensity of LPSs is overestimated (underestimated) by most models when verified against ERA-Interim (MERRA-2). The European Centre for Medium-Range Weather Forecasts and Met Office models have the best ensemble spread–error relationship for the track position and intensity, whereas the HMCR model has the worst. Most S2S models are underdispersive—more so for the intensity than the position. We find the influence of errors in the LPS simulation on the pattern of total precipitation biases in all S2S models. In most models, precipitation biases increase with forecast lead time over most of the monsoon core zone. These results demonstrate the potential for S2S models at simulating LPSs, thereby giving the possibility of improved disaster preparedness and water resources planning.

KEYWORDS: Ensembles; Forecast verification/skill; Forecasting; Hindcasts; Numerical weather prediction/forecasting; Short-range prediction

1. Introduction


Monsoon low pressure systems (LPSs) are important synoptic-scale cyclonic disturbances that are embedded in the South Asian monsoon circulation. These systems, which have a life-span of 3–5 days, most frequently form over the head of the Bay of Bengal and adjoining land area, from where they propagate in a west-northwest direction toward India (Daggupaty and Sikka 1977; Godbole 1977; Boos et al. 2015; Hunt and Parker 2016). They transport large amounts of moisture over the Indian subcontinent and are responsible for around 50% of the summer season (June–September) rainfall in India (Rastogi et al. 2018; Hunt and Fletcher 2019).

LPSs featuring surface wind speeds of 8.5–13.5 m s^{−1} or mean sea level pressure anomalies of 4–8 hPa at the center are referred to as monsoon depressions, whereas systems weaker than this are referred to as monsoon low pressure areas (Vishnu et al. 2020). Each summer, around 13–14 LPSs form, half of which intensify into monsoon depressions (Boos et al. 2015; Sandeep et al. 2018).

Since peak composite precipitation rates from LPSs are about 50 mm day^{−1} (Yoon and Chen 2005), high-impact flood events are frequent over the Indian subcontinent during the

summer. Over the last decade, LPSs have triggered at least three catastrophic floods in India: Uttarakhand (16–18 June 2013), Gujarat and Rajasthan (23–25 July 2017), and Kerala (August 2018), thereby affecting millions of people and causing unprecedented damage to property (Ray et al. 2019; Hunt and Menon 2020). Joseph et al. (2015) investigated the large-scale monsoon environment during the Uttarakhand floods in T126 (~100 km) and T382 (~38 km) model versions of the Climate Forecast System Version 2 (CFSv2) model. These two model versions could predict the occurrence of the rainfall event 10–12 days in advance, but the observed rainfall amount was underestimated by ~35%–75%, with CFS T382 outperforming CFS T126. The forecasts of heavy precipitation events can improve if models correctly predict the position and intensity of LPSs, which play such an important role in monsoon precipitation. Therefore, we would like to emphasize understanding of the predictions of LPSs on an extended time scale of 15 days given the importance of this time scale to disaster preparedness and long-term planning.

The regular occurrence of LPSs during the summer season was first reported by Eliot (1884), and since then many structural and dynamical aspects of LPSs have been explored (e.g., Godbole 1977; Ding et al. 1984; Sarkar and Choudhary 1988). In the last decade, several studies have applied feature-tracking algorithms to reanalysis datasets for identifying LPSs and examining their

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Akshay Deoras, a.s.deoras@pgr.reading.ac.uk

DOI: 10.1175/WAF-D-20-0081.1



This article is licensed under a [Creative Commons Attribution 4.0 license](http://creativecommons.org/licenses/by/4.0/) (<http://creativecommons.org/licenses/by/4.0/>).

properties (e.g., [Hurley and Boos 2015](#); [Hunt and Parker 2016](#); [Hunt et al. 2016](#); [Sørland and Sorteberg 2016](#); [Vishnu et al. 2020](#)). However, the prediction skill of LPSs remains unexplored. In contrast, there are numerous studies on the prediction skill of tropical and extratropical cyclones in various models and reanalysis datasets. [Hodges et al. \(2017\)](#) analyzed tropical cyclones (TCs) in six reanalysis datasets and inferred that low spatial resolutions of reanalyses are responsible for an underestimation of the intensity of TCs when compared to observations. [Hodges and Emerton \(2015\)](#) investigated the prediction of TCs in the Northern Hemisphere in the ECMWF ensemble and deterministic prediction systems during May–October 2008–12. They inferred that initial periods during forecasts had smaller error growth, and the location of TCs was more predictable than the intensity. They further inferred that ensemble forecasts are underdispersive (i.e., the range of ensemble forecasts is not able to fully represent all forecast states, and this might lead to observations falling outside the range of ensemble forecasts)—more for the intensity than the location. Their finding is similar to that of [Hodges and Klingaman \(2019\)](#) who inferred that the Met Office Global Forecast model is underdispersive at predicting the location of TCs in the western North Pacific. [Murakami \(2014\)](#) showed that the highest-resolution reanalyses are not always the best at simulating properties of TCs, thereby suggesting that the simulation of TCs in reanalyses is highly dependent on the model formulation and/or data assimilation.

The advent of grand ensemble prediction systems (EPSs) like the THORPEX Interactive Grand Global Ensemble (TIGGE) ([Richardson et al. 2005](#)) and the Subseasonal to Seasonal (S2S) prediction project ([Vitart et al. 2017](#)) has enabled the research community to intercompare more EPSs than in the past. [Froude \(2010, 2011\)](#) analyzed the predictions of extratropical cyclones in both hemispheres in the TIGGE dataset and concluded that the ensemble mean error of individual models of the TIGGE dataset is less than the control and ensemble members of the respective models. The better performance of the ensemble mean than ensemble members is also seen in the prediction of TCs by the ECMWF ensemble and deterministic prediction systems ([Hodges and Emerton 2015](#)). [Lee et al. \(2018\)](#) studied the prediction of TCs in six S2S models and concluded that most of these models had skill at predicting TC genesis.

Unlike TCs, LPSs spend significant duration of their lifetime over land and their propagation is confined mostly to the monsoon trough region ([Godbole 1977](#)). Thus, the results of TC predictions might not be entirely relevant for LPSs, thereby highlighting the necessity of evaluating LPS predictions. In this paper, we investigate the prediction of Indian monsoon LPSs by 11 S2S models. The prediction time scale is confined to 15 days, which is within the time scales of numerical weather prediction models. Hence, we carry out deterministic analyses of LPS predictions by considering metrics used for evaluating predictions of TCs and extratropical cyclones on similar time scales (e.g., [Froude 2010](#); [Hodges and Emerton 2015](#)). Our objective is to understand the following aspects:

- How well do S2S models represent the frequency, intensity, tracks, genesis (initial track position) and lysis (final track position) of Indian monsoon LPSs?

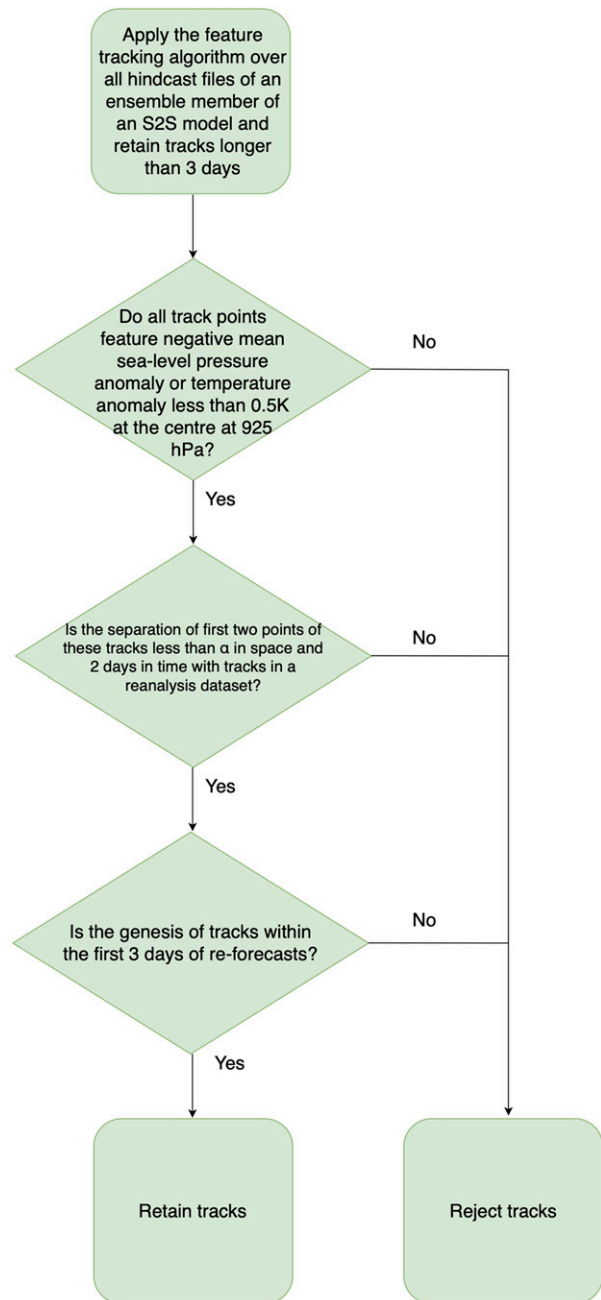


FIG. 1. Flowchart outlining the steps followed in the identification of monsoon low pressure systems in an ensemble member of an S2S model and track matching with reanalysis datasets. These steps are iterated for all ensemble members of 11 S2S models. The threshold value α is determined for each model by a sensitivity test.

- How do LPS position and intensity errors evolve with forecast lead time in S2S models?
- How statistically reliable are S2S models at predicting LPSs?
- How do forecast lead time and the presence of LPSs influence the pattern of precipitation errors in S2S models?

TABLE 1. Configuration of reforecasts in 11 S2S models, ERA-Interim, and MERRA-2 reanalysis datasets used in this paper. The intensity threshold column shows the minimum intensity of monsoon low pressure systems considered in Figs. 4–6. The minimum intensity is based on the upper-quartile values of 850-hPa relative vorticity of monsoon low pressure systems in each model, ERA-Interim, and MERRA-2.

Model	Reforecast length (days)	Resolution	Ensemble size	Reforecast frequency	Ocean coupled	Intensity threshold (10^{-5} s^{-1})
BoM	62	$\sim 2^\circ \times 2^\circ$ (T47), L17	33	Six per month	Yes	6.16
CMA	60	$\sim 1^\circ \times 1^\circ$ (T266), L40	4	Daily	Yes	4.97
ECMWF	46	$0.25^\circ \times 0.25^\circ$ (Tco639), L91: days 0–15 $0.5^\circ \times 0.5^\circ$ (Tco319), L91: after day 15	11	Two per week	Yes	4.16
ECCC	32	$0.45^\circ \times 0.45^\circ$, L40	4	Weekly	No	3.99
HMCR	61	$1.1^\circ \times 1.4^\circ$, L28	10	Weekly	No	2.97
ISAC-CNR	32	$0.8^\circ \times 0.56^\circ$, L54	5	Every 5 days	No	4.43
JMA	33	$\sim 0.5^\circ \times 0.5^\circ$ (TL479), L100	5	Three per month	No	3.92
KMA	60	$\sim 0.5^\circ \times 0.5^\circ$ (N216), L85	3	Four per month	Yes	4.31
CNRM	61	$\sim 0.7^\circ \times 0.7^\circ$ (T255), L91	15	Four per month	Yes	3.80
NCEP	44	$\sim 1^\circ \times 1^\circ$ (T126), L64	4	Daily	Yes	4.17
UKMO	60	$\sim 0.5^\circ \times 0.8^\circ$ (N216), L85	7	Four per month	Yes	4.35
ERA-Interim		$\sim 0.7^\circ \times 0.7^\circ$ (N128), L60				3.97
MERRA-2		$0.625^\circ \times 0.5^\circ$				5.11

2. Data and methods

An outline of the steps followed in the methodology is presented in Fig. 1.

a. S2S reforecasts

The S2S database consists of near-real-time forecasts (with a lag of 3 weeks) and reforecasts from 11 global operational centers: the Bureau of Meteorology, Australia (BoM), the China Meteorological Administration (CMA), the Environment and Climate Change Canada (ECCC), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Hydrometeorological Centre of Russia (HMCR), the Institute of Atmospheric Sciences and Climate of the National Research Council (ISAC-CNR), the Japan Meteorological Agency (JMA), the Korea Meteorological Administration (KMA), Météo-France/Centre National de Recherche Meteorologiques (CNRM), the National Centers for Environmental Prediction (NCEP), and the Met Office (UKMO). Each reforecast is comprised of a control reforecast and a number of perturbed reforecasts that produce ensemble members. The reforecasts are archived on a $1.5^\circ \times 1.5^\circ$ grid at a daily resolution. The S2S models have a different reforecast period, but 1999–2010 is the common reforecast period. Hence, we have considered reforecasts starting between May and September 1999–2010 in this study. Mean sea level pressure, u and v winds at 850 hPa, and temperature at 925 hPa are used for tracking and post-tracking processes. These variables are instantaneous once per day (0000 UTC). In addition, total precipitation is considered for investigating precipitation errors. The total precipitation variable is accumulated once per day for the BoM model and four times per day for other S2S models. Barring ECCC, HMCR, ISAC-CNR, and JMA models, all are ocean coupled. Details related to the configuration of reforecasts are presented in Table 1.

For this study, the following model version dates have been considered: 31 January 2017 for the JMA model, 8 June 2017

for the ISAC-CNR model, and 1 May 2014 for the CMA model. These model versions outperform the respective previous versions in terms of factors such as ensemble size. For models featuring on-the-fly configuration (i.e., reforecasts are produced at the same time as real-time forecasts) such as HMCR, ECCC, KMA, ECMWF, and UKMO, model versions used in the year 2019 have been considered for maintaining homogeneity.

b. Global precipitation measurement IMERG

The Global Precipitation Measurement (GPM) Integrated Multisatellite Retrievals for GPM (IMERG) is a merged precipitation product that provides precipitation estimates on a $0.1^\circ \times 0.1^\circ$ grid globally every half-hour (Huffman et al. 2015). The IMERG combines intercalibrated observations from satellites in the GPM constellation and is available from June 2000 in three runs: early, late, and final (Tan et al. 2019). This study uses the final runs of IMERG V06 to investigate precipitation errors, particularly over the monsoon core zone (Rajeevan et al. 2010). The performance of the IMERG V06 precipitation product has not been evaluated so far for the Indian monsoon; however, its previous versions have been intercompared (Wang et al. 2018) and compared with other datasets such as the TRMM Multisatellite Precipitation Analysis (TMPA) and IMD gauge-based dataset (Prakash et al. 2016; Liu 2016; Prakash et al. 2018) for the 2014 summer season. The IMERG shows notable improvements over TMPA in capturing heavy rainfall over India during the summer season and represents mean-monsoon rainfall more realistically. It must be noted that IMERG has difficulty in detecting rainfall over southeast and northeast India and underestimates the frequency of heavy rainfall over parts of northeast India due to the orography (Prakash et al. 2018). For this study, IMERG data has been regridded to $1^\circ \times 1^\circ$ to make a fairer comparison with the coarser S2S dataset.

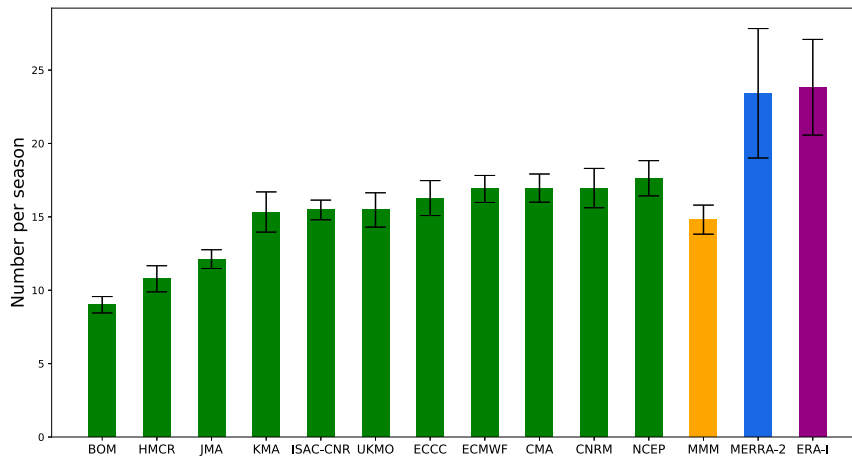


FIG. 2. Seasonal average numbers of monsoon low pressure systems in all ensemble members of 11 S2S models (green), MERRA-2 (royal blue), and ERA-Interim (purple) over the period June–September 1999–2010. The multimodel mean (MMM) is also shown (orange). Error bars show ± 1 standard deviation about the mean and calculated across years. Model results are normalized with respect to ensemble size and reforecast frequency.

c. Tracking

The identification and tracking of LPSs in this study have been performed in all ensemble members of 11 S2S models using a feature-tracking algorithm (Hunt et al. 2016, 2018) on 850-hPa relative vorticity (Deoras et al. 2021). The feature-tracking algorithm is applied to all ensemble members of 11 S2S models. The choice of using vorticity instead of mean sea level pressure for tracking is justified since the former is less sensitive to the background flow, low pressure systems are identified at an earlier stage of development and mean sea level pressure may be sensitive to the interpolation technique and representation of orography in the model (Hoskins and Hodges 2002). Moreover, production of good quality statistics is possible when vorticity is used since more features are identified (Froude 2010). The feature-tracking algorithm computes relative vorticity from 24-hourly u and v winds on the 850-hPa level in all ensemble members of 11 S2S models. To filter out small-scale vorticity features that are prevalent near orography, the spectral resolution is truncated at T63 (~ 200 km at the Equator) and all local maxima are located within a radius of 1000 km in the domain 0° – 40° N, 40° – 120° E. For each such local maximum, local positive nonzero values of relative vorticity are associated and integrated to find the centroid of relative vorticity. For each such point, the nearest neighbor is located and attached using the kd-tree nearest-neighbor algorithm (Yianilos 1993).

In the tracking algorithm, the minimum 850-hPa relative vorticity threshold was set to $1 \times 10^{-5} \text{ s}^{-1}$, which was useful in filtering out weaker eddies as suggested by Hunt et al. (2016). For further analysis, only those tracks that occurred between June and September 1999–2010, lasted for more than 3 days and had forecast lead times of less than 15 days (i.e., lysis within 15 days of each reforecasts) have been retained. Such tracks are then subjected to a post-tracking filtering process, discussed in section 2d. The feature-tracking algorithm was also

applied to the European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-I) dataset (Dee et al. 2011) and Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) (Gelaro et al. 2017). The horizontal resolution of ERA-I is approximately $0.7^\circ \times 0.7^\circ$, whereas that of MERRA-2 is $0.625^\circ \times 0.5^\circ$. The outputs for u and v winds at 850 hPa are 6-hourly in ERA-I and 3-hourly in MERRA-2. During June–September 1999–2010, 290 and 281 LPSs were identified in ERA-I and MERRA-2, respectively, which will serve as observed estimates for verifying the frequency, position and intensity of LPSs in S2S models. The additional verification against MERRA-2 will help in testing the observational uncertainty of the results. This is essential since the results might be sensitive to reanalysis datasets and verification against ERA-I alone might confer an advantage to the ECWMF model.

d. Temperature-pressure filtering

The output of the feature-tracking algorithm needs to be filtered for further diagnostics since other features such as tropical cyclones and heat lows are tracked along with LPSs. In studies related to the tracking of TCs, presence of a warm-core structure in various levels of the troposphere has been used as a criterion to segregate TCs from other tracked features (Camargo and Zebiak 2002; Camargo 2013; Camp et al. 2015). Since LPSs have a warm-over-cold core structure (Godbole 1977; Hurley and Boos 2015; Hunt et al. 2016), we focus on track filtering on the basis of temperature anomalies at the 925-hPa level. In addition, the track filtering is done using mean sea level pressure anomalies that help in removing those track points featuring nonnegative pressure anomalies. The temperature and pressure anomalies are considered at the center of the tracked system at each time step.

The anomalies in this study have been calculated by following a technique suggested by Vitart (2017). The climatologies of

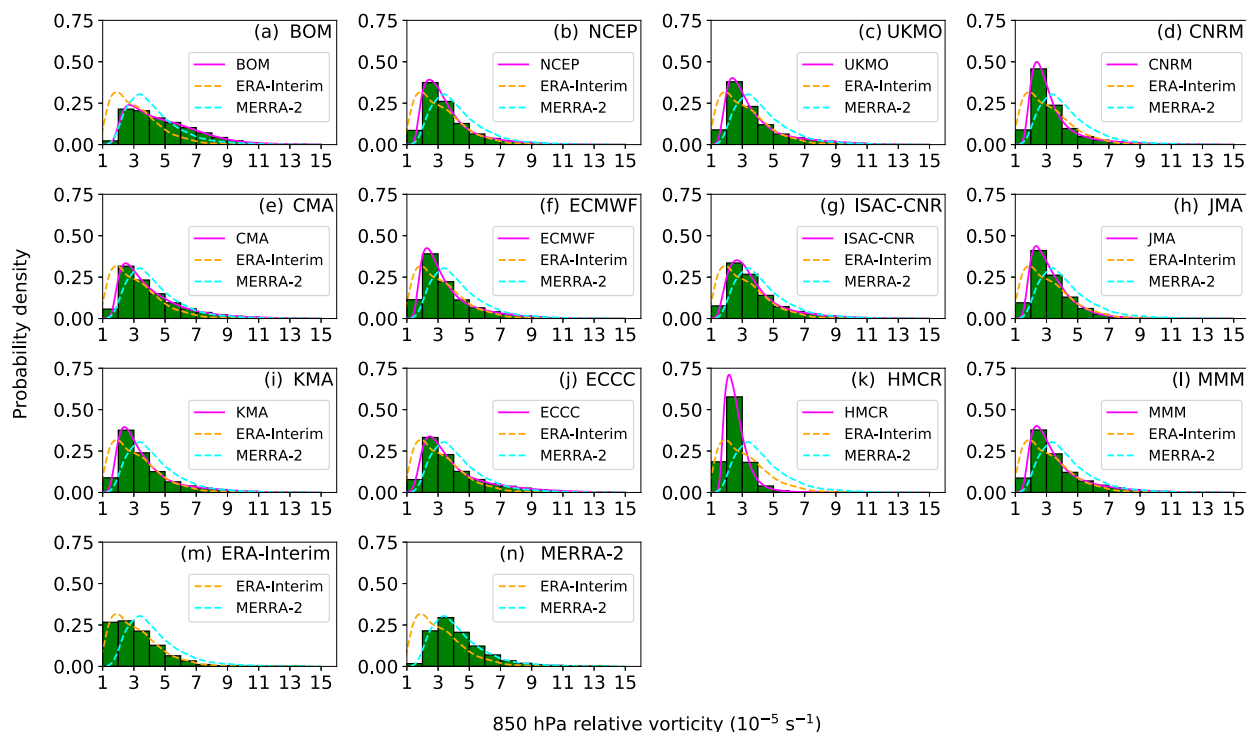


FIG. 3. Normalized histograms (green) of 850-hPa relative vorticity of monsoon low pressure systems calculated at each track point in all ensemble members of (a)–(k) 11 S2S models, (m) ERA-Interim, and (n) MERRA-2 during June–September 1999–2010. (l) The multi-model mean (MMM) is shown. Kernel density estimations using Gaussian kernels are also shown for respective individual models and MMM (solid magenta), ERA-Interim (dashed orange), and MERRA-2 (dashed cyan). Histograms are normalized with respect to ensemble size and reforecast frequency.

925-hPa temperature and mean sea level pressure have been constructed by averaging all reforecasts starting the same day and the same month, but excluding the actual year of reforecasts. For example, for a reforecast starting on 1 June 1999, the climatology will contain all reforecasts starting on 1 June 2000–10. The forecast anomalies are then calculated by subtracting the climatologies from the ensemble member. The threshold value of 925-hPa temperature anomaly was obtained from sensitivity tests conducted for all tracks in all ensemble members of 11 S2S models. For each temperature anomaly (δ) in a range, a fraction is calculated which represents a ratio between the number of track points with a temperature anomaly less than or equal to δ and the total number of track points. We then calculate the gradient of this fraction, and δ corresponding to the maximum gradient is selected as the threshold value. A similar technique was followed by Hunt and Fletcher (2019). The result of sensitivity tests suggests a threshold value of 0.5 K for all ensemble members of 11 S2S models. Thus, an entire track is removed from the tracked dataset if all of its track points have 925-hPa temperature anomaly greater than or equal to 0.5 K or nonnegative mean sea level pressure anomaly.

e. Matching methodology

To validate LPSs identified in the S2S dataset against those in reanalyses, we follow a technique of spatiotemporal matching in

which two tracks are considered to match if they meet certain predefined spatial and temporal separation criteria. Froude et al. (2007a) investigated the sensitivity of track diagnostics to the choice of spatiotemporal matching parameters in the case of extratropical cyclones. They found that the diagnostics produced from matched tracks are unaffected in spite of differences in the number of matched tracks that varied with different parametric values.

In this study, the threshold values of the spatial separation parameter were identified by sensitivity tests and the gradient technique (similar to the one discussed in section 2d) conducted for tracks in all ensemble members of 11 S2S models. Using a technique similar to Froude et al. (2007b) and Froude (2010), we consider a track in an ensemble member of an S2S model to match with a track in reanalyses if the spatial separation between the first two data points is less than a threshold value α . The values of α are 600 km for the CNRM model, 200 km for the KMA model and 500 km for the remaining S2S models. The spatial separation is considered for the first two data points instead of the entire track duration since a track in an ensemble member of an S2S model may begin very close to its corresponding track in reanalyses, but diverge with increasing forecast lead time. If multiple data points of different tracks in an ensemble member of an S2S model satisfy the spatial separation criterion for a track in reanalyses, the data point with the least temporal separation is chosen.

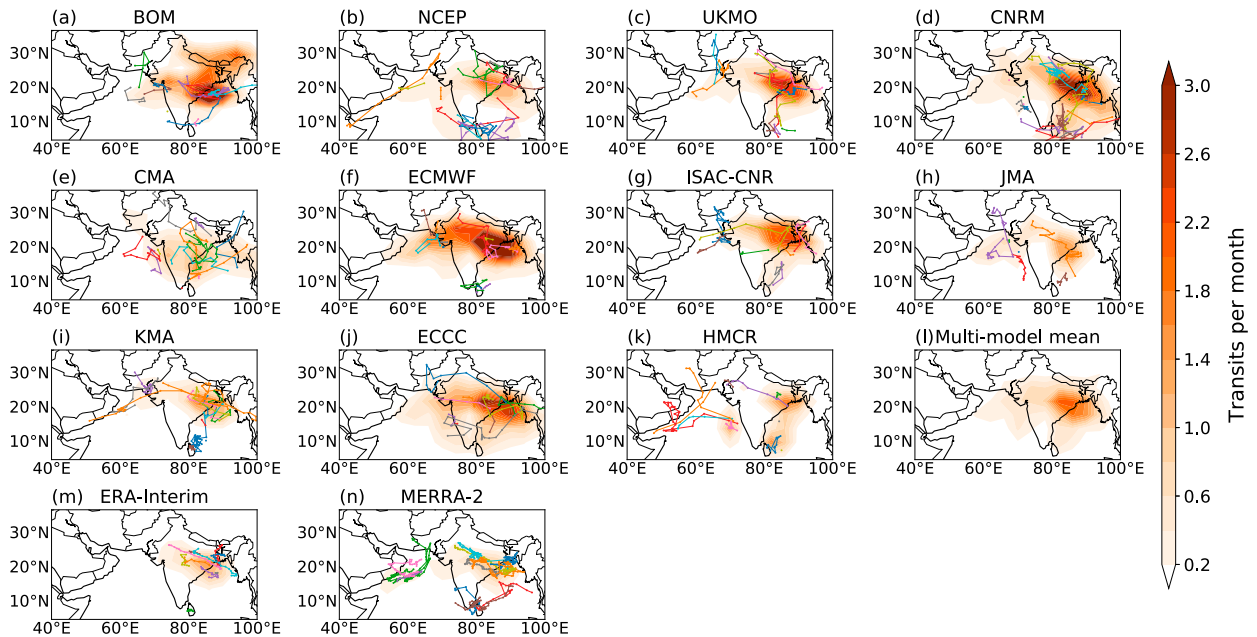


FIG. 4. Monthly mean track density (transits computed for a $4^\circ \times 4^\circ$ box centered on each grid point) of strong monsoon low pressure systems (minimum intensity equal to the upper quartile in each model) tracked in all ensemble members of (a)–(k) 11 S2S models, (m) ERA-Interim, and (n) MERRA-2 over the period June–September 1999–2010. (l) The multimodel mean track density is also shown. Solid color lines illustrate different tracks featuring within the first 15 days of a reforecast starting early June 1999 in control runs of 11 S2S models and during June–August 1999 in ERA-Interim and MERRA-2. Model results are normalized with respect to ensemble size and reforecast frequency.

For the purpose of this analysis, only those tracks that had genesis within the first three days of a reforecast or that existed already at initialization have been retained. This additional constraint helps in eliminating matches that may have occurred due to chance rather than as a real prediction (Froude et al. 2007b; Hodges and Emerton 2015).

3. Climatology of LPSs

In this section, we present the following verification results related to LPSs:

- Seasonal average numbers.
- Intensity distribution.
- Track, genesis, and lysis density.

a. Seasonal numbers

The seasonal average numbers of LPSs in all ensemble members of 11 S2S models, MERRA-2 and ERA-I during June–September 1999–2010 are shown in Fig. 2 along with the multimodel mean, for forecast lead times of less than 15 days. The S2S models exhibit a prominent spread in the simulated frequency of LPSs, ranging from 9 (± 0.56) in the BoM model to 18 (± 1.20) in the NCEP model. Compared to 23.83 (± 3.26) and 23.42 (± 4.41) LPSs simulated per season by ERA-I and MERRA-2, respectively, all S2S models underestimate the frequency, with only 14.81 (± 0.99) LPSs per season in the multimodel mean. The range in parentheses indicates one standard deviation

about seasonal average numbers of LPSs calculated across 1999–2010.

The low-frequency of LPSs simulated by models such as the BoM could be related to a weak and poorly defined monsoon trough, which provides cyclonic vorticity in the lower troposphere to spin up LPSs (Godbole 1977). In addition, the frequency is also dependent on intraseasonal oscillations such as the boreal summer intraseasonal oscillation (BSISO; Kikuchi and Wang 2010). These aspects will be examined in a future study.

b. Intensity distribution

The probability density of intensity (850-hPa relative vorticity) of LPSs in all ensemble members of 11 S2S models, multimodel mean, ERA-I, and MERRA-2 is shown in Fig. 3. The intensity is considered at the center of each track at each lead time (up to 15 days) since the maximum relative vorticity is observed at the center of LPSs in the lower troposphere (Godbole 1977). Gaussian kernel density estimation, which is a nonparametric way to estimate the probability density function using Gaussian kernels (Scott 2015), is also shown for comparison. Differences in the intensity distribution of LPSs can be found among different S2S models; however, in all these models and the multimodel mean (Fig. 3l), the largest probability density is observed for intensity in the range $2\text{--}3 \times 10^{-5} \text{ s}^{-1}$, which is in agreement with ERA-I (Fig. 3m). This result was anticipated since not all LPSs intensify into stronger systems such as monsoon depressions. For all S2S models except the BoM model, the probability density of track points

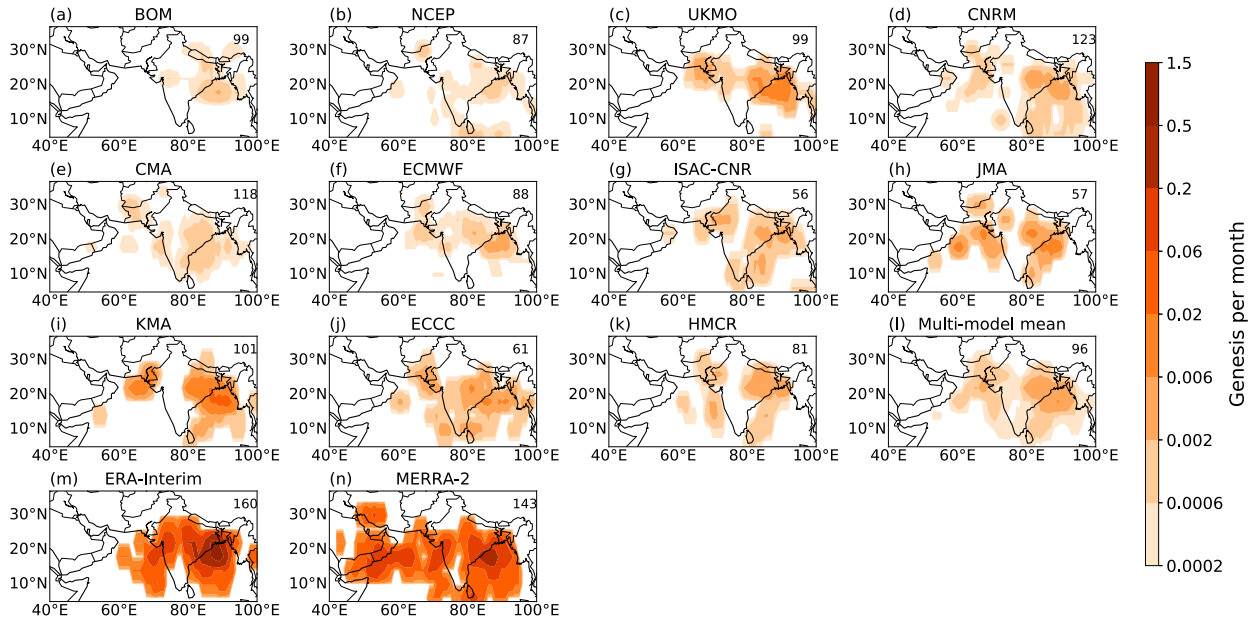


FIG. 5. As in Fig. 4, but for genesis density. The number of genesis points is shown in each subplot.

featuring intensity more than $3 \times 10^{-5} \text{ s}^{-1}$ decreases rapidly, which is also seen in ERA-I, but not in MERRA-2. The largest probability density in MERRA-2 is seen for the intensity in the range $3\text{--}4 \times 10^{-5} \text{ s}^{-1}$, following which there is a rapid decline in the probability density. For the BoM and CMA models (Figs. 3a,e, respectively), the probability density of track points featuring intensity greater than or equal to $6 \times 10^{-5} \text{ s}^{-1}$ is larger than ERA-I, but equal to MERRA-2 up to $8 \times 10^{-5} \text{ s}^{-1}$. It must be noted that unlike ERA-I, all S2S models and MERRA-2 have a noticeably smaller probability density of

track points featuring intensity in the range $1\text{--}2 \times 10^{-5} \text{ s}^{-1}$ than $2\text{--}3 \times 10^{-5} \text{ s}^{-1}$. This variation is due to a greater genesis of weaker LPSs in ERA-I than in S2S models and MERRA-2. In addition, LPSs in ERA-I have shorter lifetime than in S2S models and MERRA-2.

c. Track density

Among monsoon low pressure areas and strong LPSs (SLPSs) such as monsoon depressions, the latter are known to have produced more catastrophic impacts in the Indian

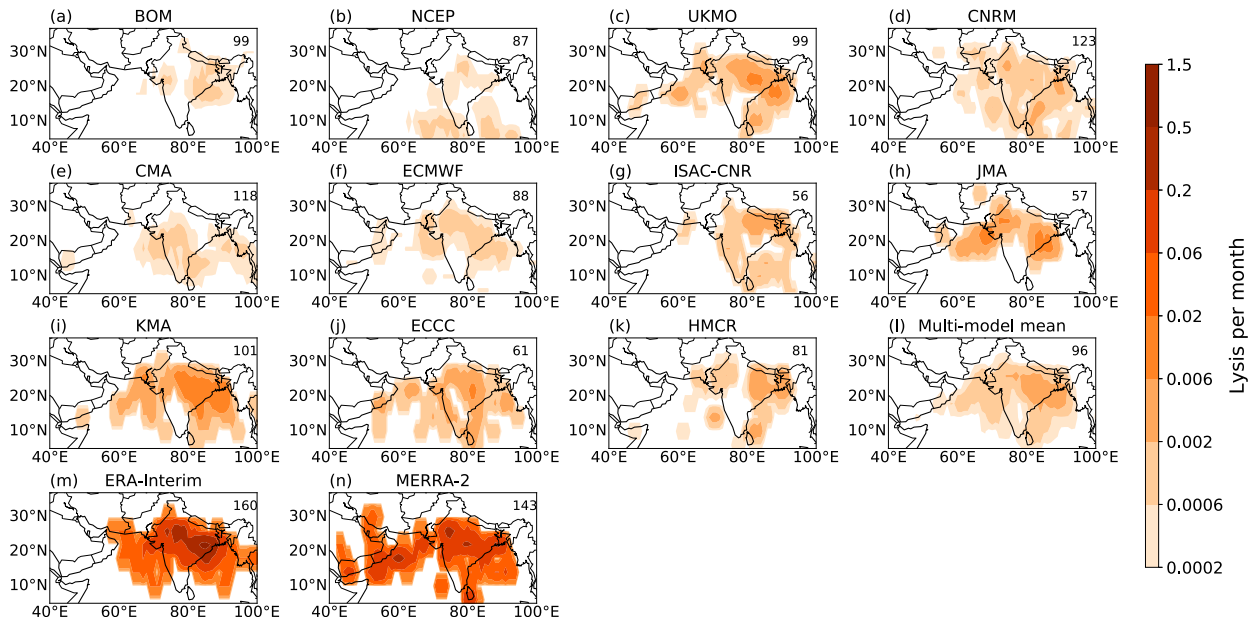


FIG. 6. As in Fig. 4, but for lysis density. The number of lysis points is shown in each subplot.

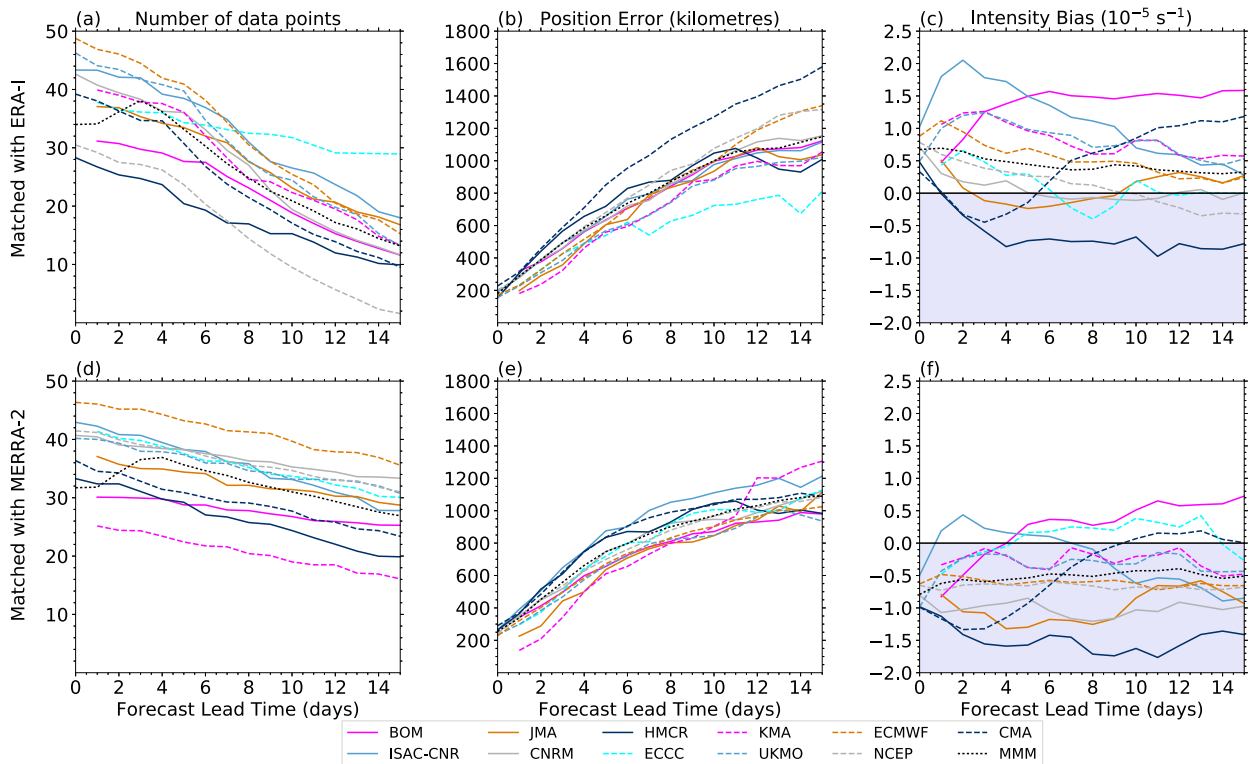


FIG. 7. (a) The number of data points of tracks matched with ERA-I that are included in the statistics for (b), (c), and Figs. 9 and 10 as a function of forecast lead time (days). The results are normalized with respect to ensemble size and reforecast frequency. (b) Error in the position of monsoon low pressure systems (km) as a function of forecast lead time (days); and (c) bias in the intensity (10^{-5} s^{-1}) of monsoon low pressure systems as a function of forecast lead time (days). The shaded region indicates negative bias in the intensity. These results are calculated for all ensemble members of 11 S2S models. The multimodel mean is also shown in dotted black in each subplot. (d)–(f) As in (a)–(c), but for tracks matched with MERRA-2. Output from step 0 is not available for the BoM, JMA, KMA, and ECCC models in all subplots.

subcontinent than the former. Thus, it is essential to understand how S2S models represent transits of SLPs and how they differ from reanalyses. In this study, we define SLPs as LPSs with minimum 850-hPa relative vorticity greater than or equal to the third quartile of 850-hPa relative vorticity of all LPSs in an individual S2S model (or reanalysis). The threshold intensity values for S2S models, MERRA-2 and ERA-I are provided in Table 1. The track density of SLPs is calculated by binning tracks in $4^\circ \times 4^\circ$ boxes and then normalizing with respect to ensemble size and reforecast frequency, similar to the methods used in Camp et al. (2015) and Lee et al. (2018). Example tracks starting early June 1999 in the control runs are shown in the respective S2S models (Figs. 4a–k) and during June–August 1999 in ERA-I (Fig. 4m) and MERRA-2 (Fig. 4n).

All 11 S2S models are capable of simulating tracks of SLPs over the head of the Bay of Bengal and adjoining land area (Figs. 4a–k); the NCEP, UKMO, CNRM, ECMWF, and KMA models perform better than other S2S models. Tracks over these regions are also observed in the multimodel mean (Fig. 4l), which is in agreement with ERA-I (Fig. 4m) and MERRA-2 (Fig. 4n). In the BoM (Fig. 4a) and JMA (Fig. 4h) models, SLPs tracks occur further south than in ERA-I,

whereas in CMA, the track direction is westward as a result of easterly midtropospheric steering winds over the head of the Bay of Bengal and central India (not shown). The CMA and JMA models have a smaller track density compared to other S2S models and reanalyses. In the HMCR model, tracks over west-central India are not observed due to faster lysis (weaker intensity) of SLPs and their low count. Thus, S2S models exhibit regional biases in simulating tracks of SLPs, but the performance of the MMM is good in general.

d. Genesis and lysis

In this subsection, genesis and lysis locations of SLPs are examined. Figures 5 and 6 show genesis and lysis densities of SLPs, which have been calculated by following the same process discussed in section 3c. In addition, the density function is sampled by centering densities on each grid point separated by 1° . This process is essential since low densities and large domain size lead to numerical artifacts in contours. Most S2S models correctly represent the primary genesis region over the head of the Bay of Bengal and adjoining land area, which is also represented in the multimodel mean (Fig. 5l). A secondary genesis region over the eastern Arabian Sea and the western coast of India is visible in the multimodel mean (Fig. 5l), ERA-I

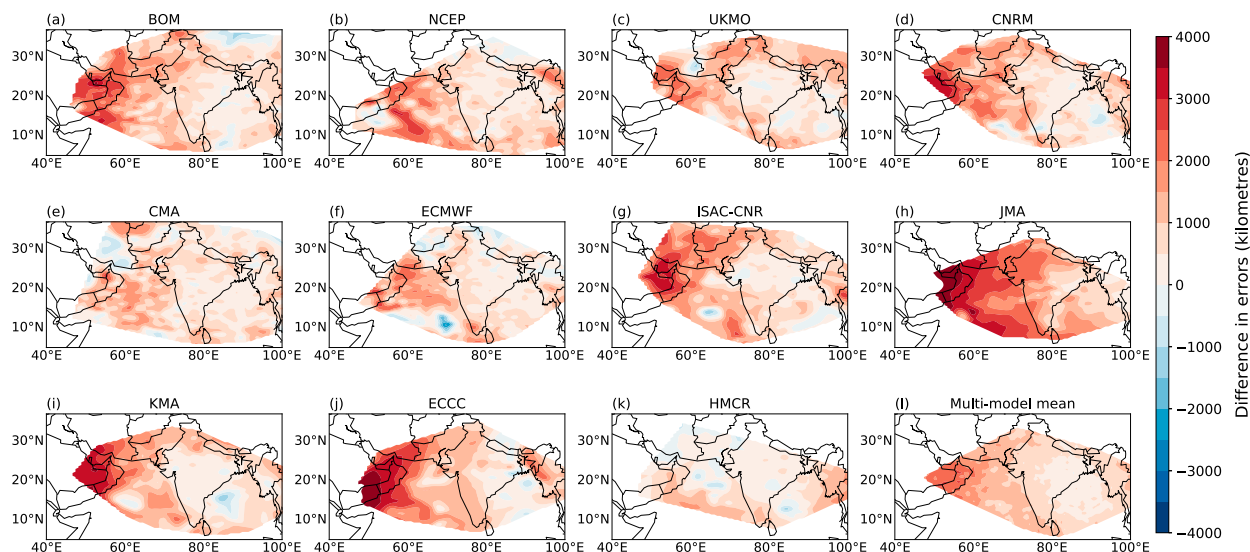


FIG. 8. (a)–(k) Difference in the position error (km) of monsoon low pressure systems between lead times 0–3 and 12–15 days in 11 S2S models. The difference is calculated by subtracting the position errors during 0–3-day lead times from 12- to 15-day lead times for tracks matched with ERA-I. (l) The multimodel mean of the difference in the position error is also shown.

(Fig. 5m), and MERRA-2 (Fig. 5n); however, genesis over this region is not represented in models such as the BoM (Fig. 5a).

In terms of lysis, most S2S models including the multimodel mean (Fig. 6l) represent the primary lysis region over eastern India and the secondary lysis region over parts of western and central India; the UKMO, ECMWF, JMA, and KMA models have representations that are the most similar to ERA-I and MERRA-2. It must be noted that there are fewer genesis and lysis points in 11 S2S models than in ERA-I and MERRA-2, which could be due to factors such as differences in the intensity of the monsoon trough. However, a detailed investigation of such factors is beyond the scope of this study. Thus, S2S models do a good job in general at simulating the primary genesis, primary lysis and secondary lysis regions. The MMM outperforms individual S2S models since it correctly simulates all genesis and lysis locations.

4. The skill of LPS predictions

In this section, we discuss the following results:

- Relative skill of ensemble members.
- Spatial distribution of position errors.
- Control and ensemble mean error.
- Ensemble spread–error relationship.

a. Relative skill of ensemble members

In this subsection, the relative skill of all ensemble members of 11 S2S models at predicting the position and intensity of LPSs is examined. To calculate the position error, the geodesic separation distance between each pair of matched tracks at each lead time is considered during lead times of 0–15 days.

Figures 7a and 7d show the number of track points in S2S models that match with those in ERA-I and MERRA-2,

respectively. These track points have been included in the statistics discussed in this subsection as well as sections 4c and 4d. The multimodel mean is also shown. The results are normalized with respect to ensemble size and reforecast frequency. The number of data points decreases with an increase in lead time due to the lysis of LPSs. This decrease is rapid in many models after 4 days since only those LPSs that had their genesis within the first 3 days of reforecasts or which existed already at initialization have been considered (see section 2e).

Figures 7b and 7e show the position error of LPSs for all ensemble members of 11 S2S models when matched with ERA-I and MERRA-2, respectively, whereas Fig. A1 shows these position errors for each model separately when LPSs are matched with ERA-I. It can be found that the position error increases with lead time in all models as well as the multimodel mean, and this result is independent of the choice of the reanalysis dataset. In addition, there are differences in the skill of S2S models. When ERA-I is used for verification (Fig. 7b), the CMA model has the lowest skill (the largest position error) for all lead times, whereas the ECCC, UKMO, JMA, and KMA models have higher skill (smaller position error) than most S2S models. The CMA model has ~3 days less skill than among the best performing models such as UKMO. At 6 days, the position error in the CMA model is 1000 km that becomes ~1600 km by 15 days lead time. The large position error in this model can be understood from the bias that leads to the westward propagation of LPSs (such as SLPs in Fig. 4e) instead of the observed west-northwest propagation in ERA-I (Fig. 4m) and MERRA-2 (Fig. 4n). The small position errors in the UKMO, JMA and KMA models and the large position error in the CMA model are also seen when LPSs are matched with MERRA-2 (Fig. 7e). However, the magnitude of the error in the CMA model is smaller as a result of a greater number of westward moving LPSs in MERRA-2 than in ERA-I. It must

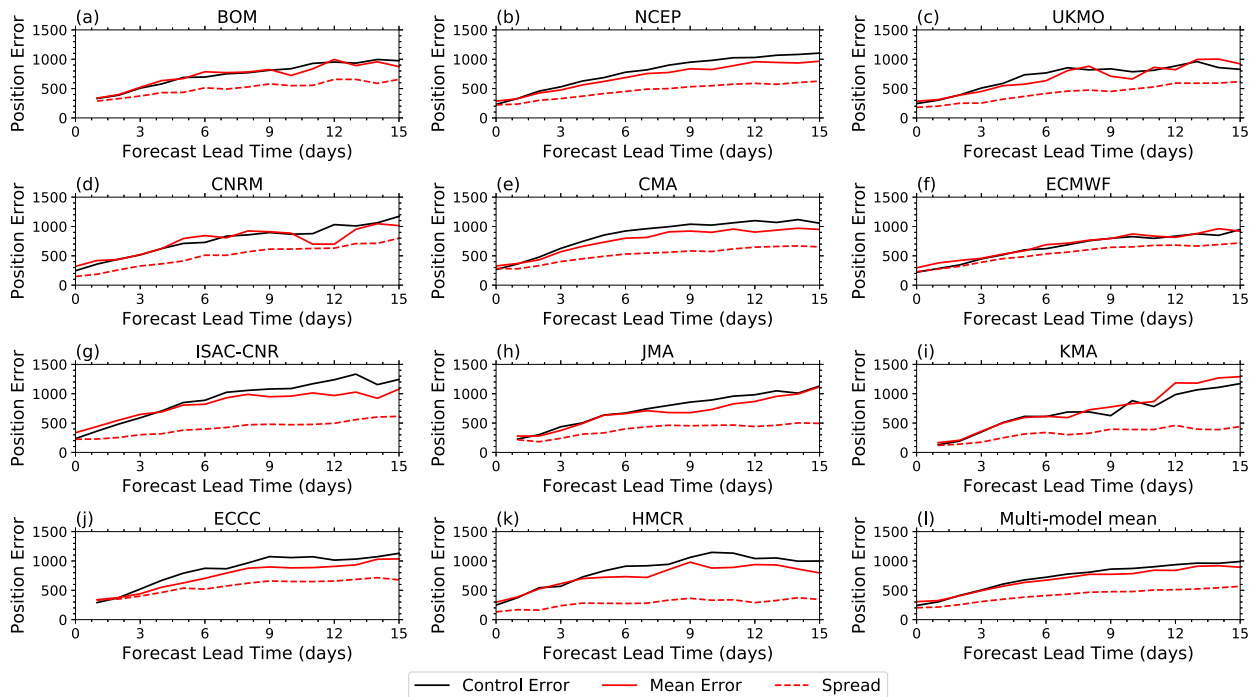


FIG. 9. Ensemble mean error (solid red), control forecast error (solid black), and spread (dashed red) in the position of monsoon low pressure systems in (a)–(k) 11 S2S models and (l) the multimodel mean. Errors and spread are calculated with respect to ERA-Interim. The unit of position error is kilometers. Output from step 0 is not available for the BoM, JMA, KMA, and ECCC models.

be noted that the range of the position error in the MME is very similar in both verification results.

Figures 7c and 7f show biases in the intensity of LPSs for all ensemble members of S2S models when matched with ERA-I and MERRA-2, respectively, whereas Fig. A2 shows biases for each model separately when LPSs are matched with ERA-I. Similar to the position error, differences can be seen in the skill of S2S models; when LPSs are matched with ERA-I, many models including the multimodel mean overestimate the intensity of LPSs at all lead times, except for the HMCR model that underestimates the intensity beyond lead times of 1 day. The intensity bias is the smallest at most lead times for the JMA and CNRM models, whereas models such as the BoM and HMCR exhibit the largest bias. A rapid increase in the bias can be observed in the HMCR, BoM and ISAC-CNR models at shorter lead times. However, when LPSs are matched with MERRA-2, most models including the multimodel mean underestimate the intensity of LPSs at all lead times, except for models such as the BoM and ISAC-CNR. This underestimation is a consequence of stronger LPSs in MERRA-2 than in ERA-I. It must be noted that the HMCR (BoM) model exhibits the largest negative (positive) intensity bias in both verification results and the overall pattern of biases among most models show consistency. These results suggest that using ERA-I for verification does not give an advantage to the ECMWF model since the latter does not exhibit the smallest position error and intensity bias when verified against both reanalysis datasets. This is opposite to the findings of Froude (2010, 2011) in which verification of the results against ECMWF

analysis was considered to be a reason for the best performance of the TIGGE-ECMWF model. The bias cannot be calculated for the position error since this error is positive.

b. Spatial distribution of position errors

In this subsection, we investigate how forecast lead time influences the spatial distribution of position errors of LPSs in 11 S2S models. Figure 8 shows the difference in position errors of LPSs that match with those in ERA-I between lead times 0–3 and 12–15 days in 11 S2S models, in order to quantify how the errors may have changed over time. The multimodel mean is also shown (Fig. 8l). The difference is calculated by subtracting position errors during 0–3-day lead times from 12- to 15-day lead times. The results confirm that position errors increase with forecast lead time over most of the domain in all S2S models, which is in agreement with Fig. 7b. The HMCR model (Fig. 8k) outperforms the multimodel mean (Fig. 8l) for difference in the position error. In all models and the multimodel mean, position errors are larger over the Arabian Sea than over the Bay of Bengal at lead times of more than 12 days. However, large position errors over the Arabian Sea are not seen when MERRA-2 is considered for verification (not shown). A greater number of LPSs, which have their genesis over the Arabian Sea, are tracked in MERRA-2 than in ERA-I. In addition, LPSs reaching the Arabian Sea from the Bay of Bengal have a longer life-span (and thus persist to longer lead times) than those having their genesis over the Arabian Sea. These factors reduce the position error over the Arabian Sea in MERRA-2.

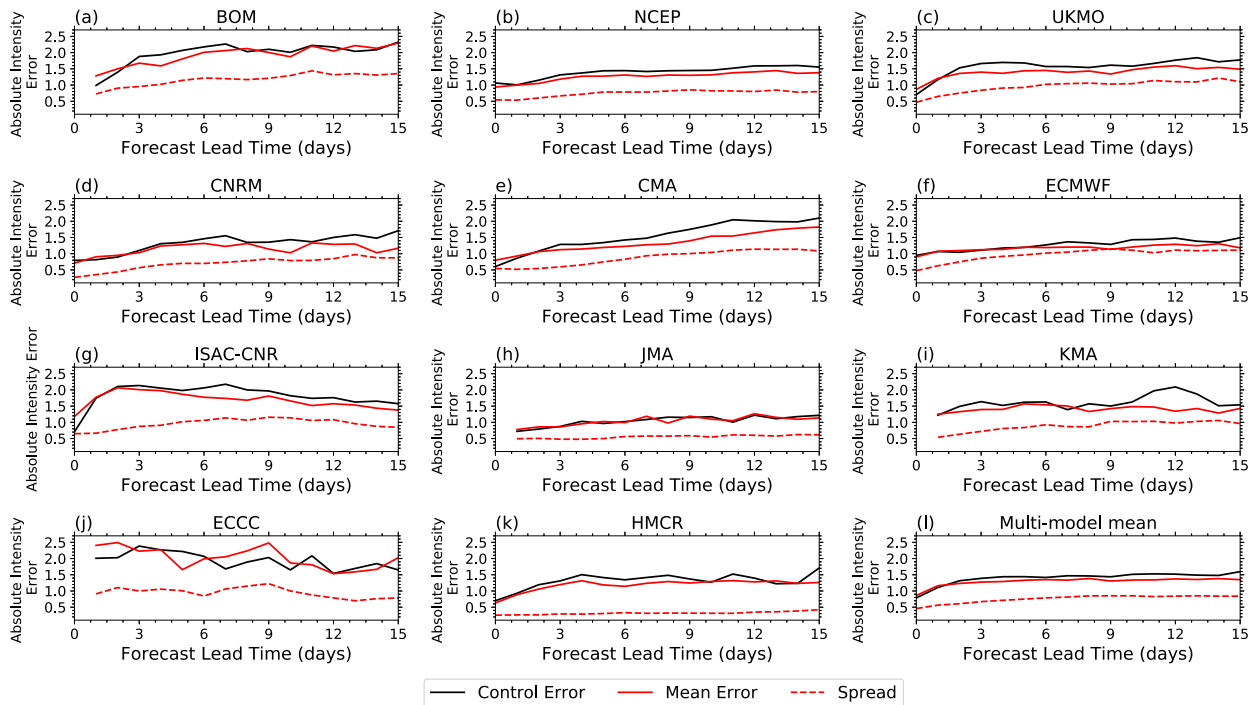


FIG. 10. Ensemble mean error (solid red), control forecast error (solid black), and spread (dashed red) in the intensity of monsoon low pressure systems in (a)–(k) 11 S2S models and (l) the multimodel mean. Errors and spread are calculated with respect to ERA-I. The unit of intensity error is 10^{-5} s^{-1} . Output from step 0 is not available for the BoM, JMA, KMA, and ECCC models.

c. Control and ensemble mean error

An important advantage of an ensemble prediction system (EPS) is that an ensemble mean provides a superior forecast compared to a control since the process of averaging removes the less predictable spatial scales (Leith 1974; Toth and Kalnay 1993, 1997; Hodges and Emerton 2015). In this subsection, the skill of the ensemble mean at predicting the position and intensity of LPSs is compared with the control forecast for 11 S2S models. The ensemble mean error for an EPS is calculated by first computing ensemble mean tracks of all LPSs in ensemble members that match LPSs in ERA-I. For an EPS, the number of ensemble members that have matching tracks and the length of such tracks in each ensemble member varies for different LPSs. In previous studies examining the prediction skill of EPSs, the tracks considered were those that were present in at least five ensemble members (Froude et al. 2007b; Froude 2010, 2011). Since several S2S models have at most five ensemble members (see Table 1), we consider only those tracks in this diagnostic that are present in at least two ensemble members. For an LPS, the mean position error in an EPS is calculated as the mean geodesic separation distance between the ensemble mean track and its corresponding matched track in ERA-I at each lead time. This process is iterated for all LPSs in 11 S2S models to obtain ensemble mean errors for all S2S models. It must be noted that for the BoM model, the mean of the control errors is considered since the EPS consists of three model versions and thus three control runs.

Figures 9 and 10 show ensemble mean error, control error and ensemble spread in LPS position and intensity, respectively, in all

S2S models when LPSs are matched with those in ERA-I. Similar analyses are carried out using MERRA-2 (not shown). The ensemble spread will be discussed in section 4d. For the position error, the ensemble mean provides an advantage over the control (i.e., the ensemble mean error is less than the control error) for most S2S models; however, it provides very little advantage in the KMA model (Fig. 9i) at lead times greater than 10 days. These results are similar when MERRA-2 is used for verification—the ensemble mean in most models provides an advantage over the control run, but the difference between them is smaller than in ERA-I.

For the LPS intensity, the ensemble mean provides a little advantage over the control run for some S2S models such as the NCEP (Fig. 10b), UKMO (Fig. 10c), CMA (Fig. 10e) and ECMWF (Fig. 10f) when ERA-I is considered. It does not provide any distinct advantage for the JMA model (Fig. 10h). However, the ensemble mean in most models provides a greater advantage over the control run when MERRA-2 is considered for verification instead of ERA-I. Thus, the ensemble mean is more advantageous over the control forecast for the intensity of LPSs than their position when MERRA-2 is considered for verification. This result agrees with the findings of Froude (2010, 2011) for extratropical cyclones in the TIGGE dataset. It must be noted that the multimodel mean provides an advantage over the multimodel control for the position and intensity of LPSs when verified against ERA-I and MERRA-2.

d. Ensemble spread–error relationship

To ascertain the reliability of S2S models at predicting the position and intensity of LPSs, the ensemble spread–error

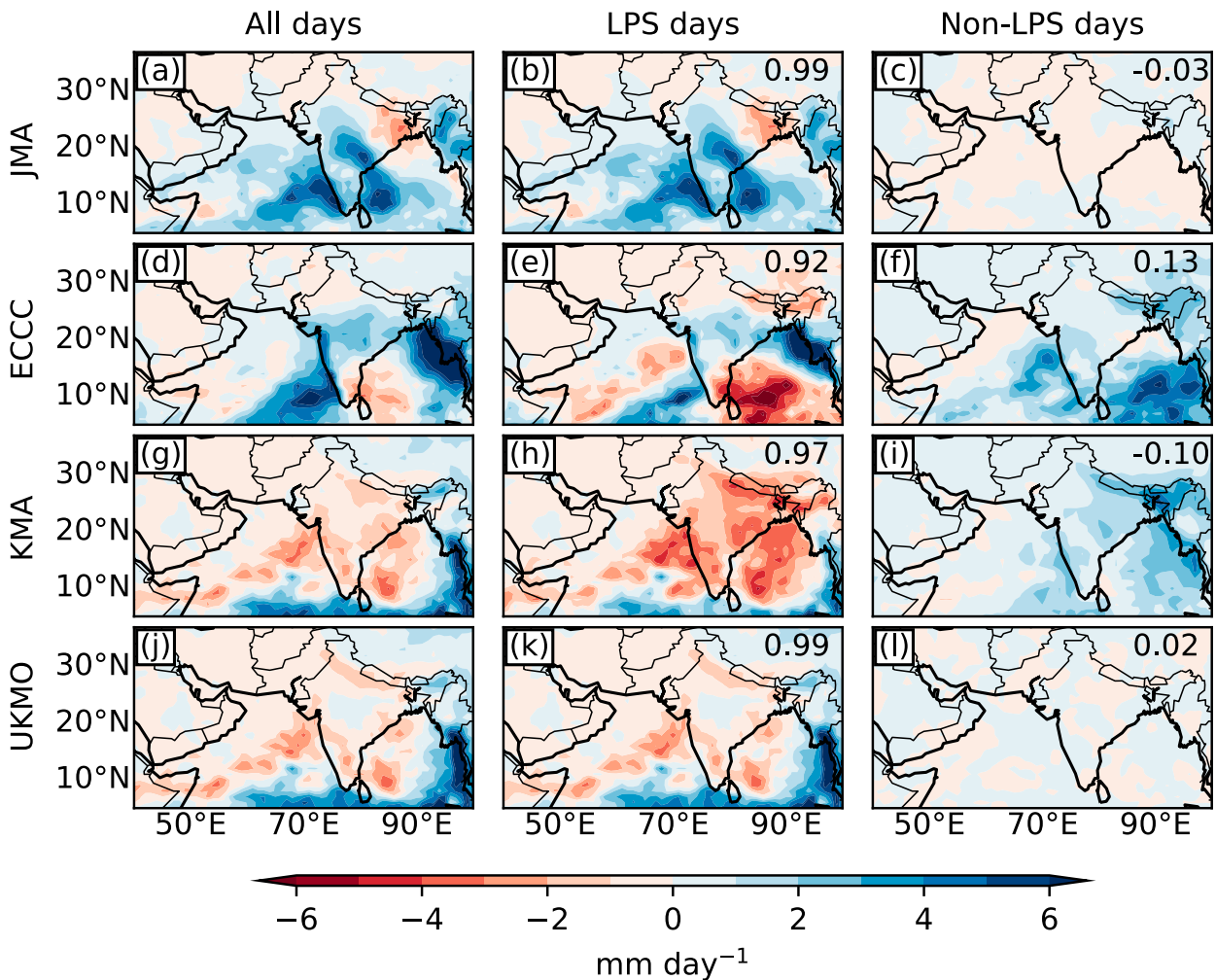


FIG. 11. Difference in daily precipitation (mm day^{-1}) in the (a)–(c) JMA, (d)–(f) ECCC, (g)–(i) KMA, and (j)–(l) UKMO models. For every model, the difference is calculated as model precipitation minus GPM IMERG precipitation for (left) all days (days 0–3 and 12–15), (center) low pressure system (LPS) days, and (right) non-LPS days in the same time range. Numbers indicate pattern correlation coefficient between LPS days and all days and non-LPS days and all days.

relationship is investigated. For a statistically reliable EPS, the ensemble spread should be equal to the ensemble mean error (Froude 2010). This means that the ensemble spread should be able to cover all possible forecast outcomes and predict the forecast error (Leutbecher and Palmer 2008; Hopson 2014). However, EPSs tend to display underdispersion since not all sources of forecast uncertainties related to initial conditions and model errors are simulated (Buizza et al. 2005).

In this study, the ensemble spread for an LPS in an S2S model is calculated as the mean geodesic separation distance between the ensemble mean track and corresponding ensemble member tracks at each lead time. The ensemble spread for all S2S models is then calculated by repeating the process for all matched LPSs in all S2S models. For the position of LPSs (Fig. 9), the BoM (Fig. 9a), NCEP (Fig. 9b), UKMO (Fig. 9c) and ECMWF (Fig. 9f) models have the best ensemble spread–error relationship (i.e., the curves showing the ensemble spread and ensemble mean error are the closest to each other) when

ERA-I is used for verification. The other S2S models are underdispersive to varying degrees, with the HMCR model having the worst ensemble spread–error relationship. In MERRA-2 (not shown), the ECMWF model has the best ensemble spread–error relationship among all S2S models; this suggests that the result is not sensitive to the reanalysis dataset used for verification.

For the intensity of LPSs (Fig. 10), there are larger differences between ensemble mean error and ensemble spread than the position of LPSs. The NCEP (Fig. 10b), UKMO (Fig. 10c) and ECMWF (Fig. 10f) models have the best ensemble spread–error relationship, whereas the HMCR model (Fig. 10k) has the worst. These results are consistent when verified against MERRA-2 (not shown). The ensemble spread depends on the number of ensemble members and the perturbation method. Despite having fewer ensemble members, many S2S models have better ensemble spread–error relationships than the HMCR model. This suggests that the reason for the worst

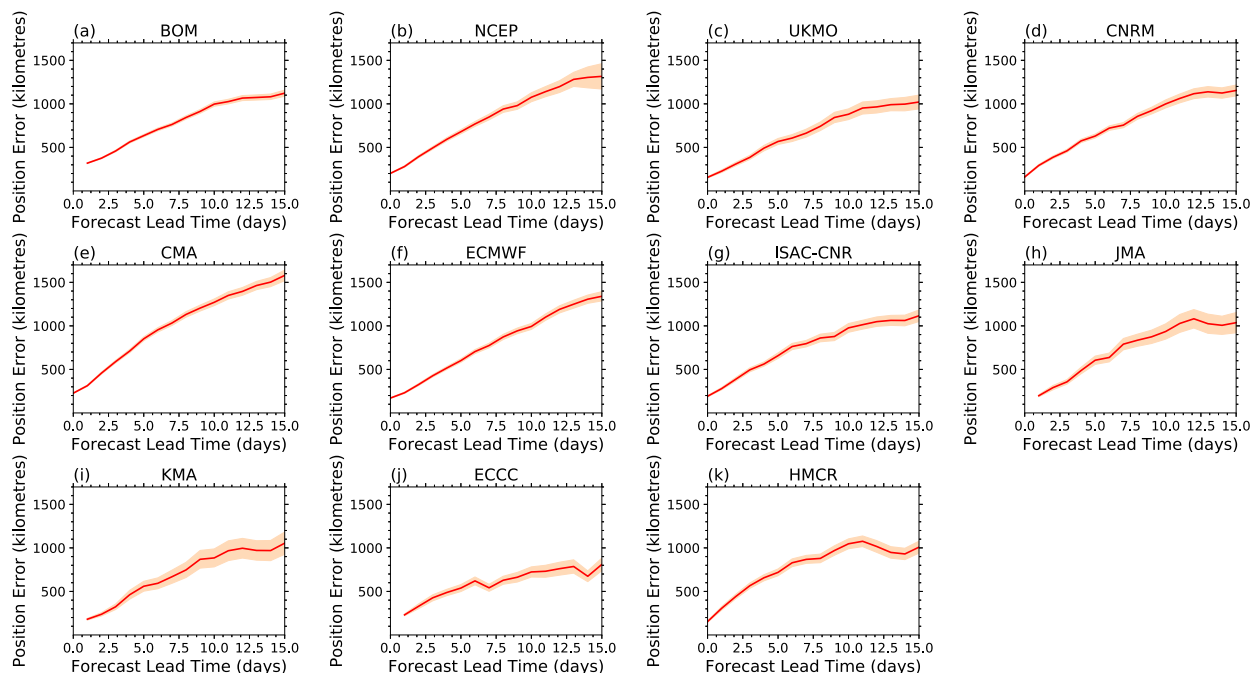


FIG. A1. (a)–(k) As in Fig. 7b, but position errors are shown separately for 11 S2S models. The transparent shaded regions indicate the 95% confidence intervals for the mean position errors, which are computed from the standard errors. Output from step 0 is not available for the BoM, JMA, KMA, and ECCC models.

ensemble spread–error relationship in this model is perhaps the perturbation method. However, this analysis requires a sensitivity test with the model, which is outside the scope of our study. Compared to the position of LPSs, models are more under-dispersive for the intensity; this result is similar to that of extratropical cyclones in the TIGGE dataset (Froude 2010, 2011).

5. Precipitation errors

In this section, we investigate how forecast lead time and the presence or absence of LPSs influence seasonal mean precipitation errors in 11 S2S models. The S2S precipitation data has been regridded to $1^\circ \times 1^\circ$. The difference in daily precipitation is calculated by subtracting IMERG precipitation from S2S precipitation for forecast lead times of 12–15 days minus 0–3 days. This difference is calculated for three cases: all days in the time range, LPS days (when an LPS was present in the domain) and non-LPS days. The pattern correlation coefficient is also calculated to evaluate the strength of the relationship between LPS days and all days as well as non-LPS days and all days.

Figure 11 shows differences in daily precipitation in the JMA, ECCC, KMA and UKMO models. In the JMA (Fig. 11a) and ECCC (Fig. 11d) models, wet biases of $2\text{--}3\text{ mm day}^{-1}$ are visible over most of the monsoon core zone, which increase to $\sim 4\text{ mm day}^{-1}$ over western India. However, a mostly dry bias is visible in the same regions in the KMA (Fig. 11g) and UKMO (Fig. 11j) models, which have a peak value of $\sim -3\text{ mm day}^{-1}$. The precipitation difference for other S2S models is shown in the appendix (Figs. A3 and A4). Excluding the ISAC-CNR

model, other models exhibit wet biases over most of the monsoon core zone; these biases are as large as $\sim 20\text{ mm day}^{-1}$ in the CMA model. Dry biases in the KMA and ISAC-CNR models could be due to moisture biases, but this cannot be investigated due to the unavailability of moisture-related parameters in the output data of these models. These results suggest that precipitation error over the monsoon core zone increases with forecast lead time in all 11 S2S models except the KMA, UKMO, and ISAC-CNR models. The multimodel mean of Coupled Model Intercomparison Project-5 (CMIP5) and CMIP3 models exhibit wet biases (dry biases) over eastern parts of the Arabian Sea (monsoon core zone) during the summer season (Sperber et al. 2013); similar wet (dry) biases are found in the JMA, ECCC, and CMA (ISAC-CNR and KMA) models in this study.

The strong wet bias along the western coast of India in models such as CMA is due the intensification of an offshore trough (Francis and Gadgil 2006) at 12–15-day lead times compared to 0–3 days. Over parts of the western coast, MSLP decreases by $\sim 2\text{ hPa}$ and specific humidity increases by $\sim 3 \times 10^{-3}\text{ kg kg}^{-1}$ at the 850-hPa level during 12–15-day lead times (not shown). In the UKMO model (not shown), specific humidity over the same region decreases during 12–15-day lead times, which causes the dry bias. The pattern correlation coefficient between LPS days and all days is 0.99 in most S2S models, suggesting that LPSs influence the pattern of precipitation errors in S2S models. Even in the ECCC and KMA models, the pattern of precipitation errors is the most similar between all days and LPS days instead of all days and non-LPS days. It must be noted that the precipitation difference for 11

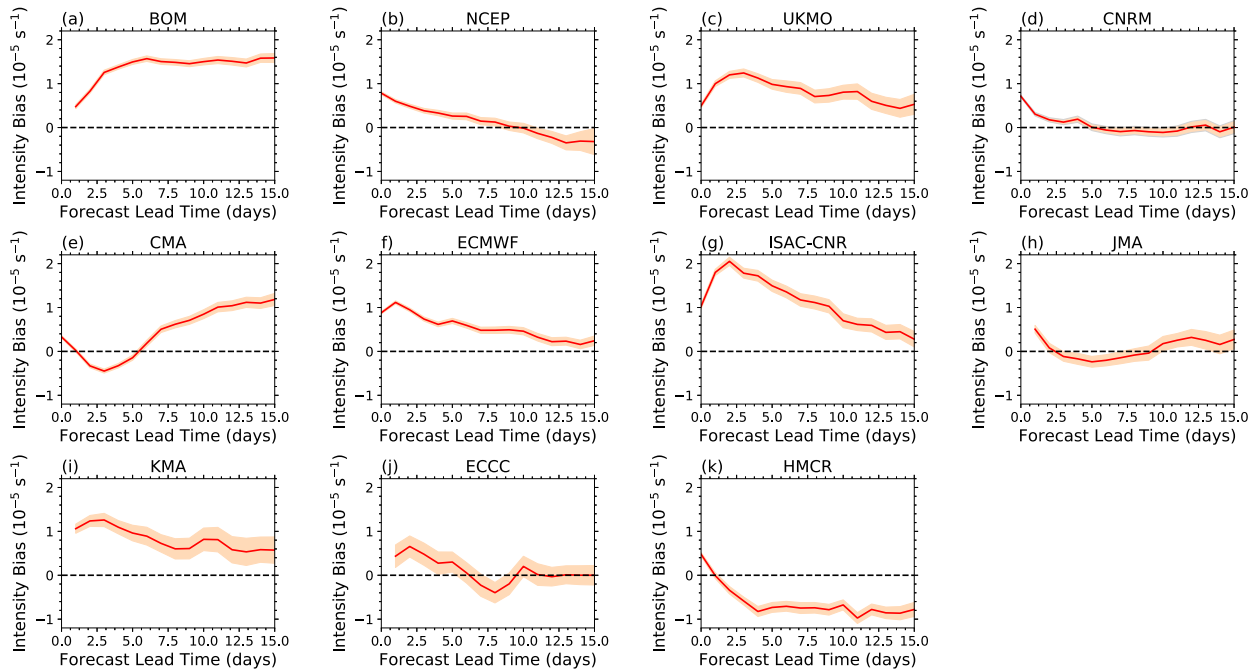


FIG. A2. (a)–(k) As in Fig. 7c, but intensity biases are shown separately for 11 S2S models. The transparent shaded regions indicate the 95% confidence intervals for the mean intensity errors, which are computed from the standard errors. Output from step 0 is not available for the BoM, JMA, KMA, and ECCC models.

S2S models is similar even when matched LPSs are considered instead of unmatched LPSs (not shown).

6. Discussion and conclusions

In this paper, we have analyzed the prediction of Indian monsoon low pressure systems (LPSs) by 11 models of the Subseasonal-to-Seasonal (S2S) prediction project (Vitart et al. 2017). LPSs are a crucial component of the Indian monsoon since they produce substantial rainfall in the Indian subcontinent during the summer season. In spite of their important role for water supply and for triggering catastrophic flood events in the subcontinent, examining the potential for their predictability on the time scales of numerical weather prediction and extended range models remains less explored than for other phenomena such as tropical cyclones. We used a feature-tracking algorithm to track LPSs in all ensemble members of 11 S2S models during the common reforecast period of June–September 1999–2010. Tracks were then subjected to a post-tracking filtering process in which tropical cyclones and heat lows were eliminated. The retained LPSs were compared with 290 and 281 LPSs identified in ERA-Interim reanalysis (ERA-I) and Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) datasets, respectively, for the purpose of verification. The results can be summarized as follows:

a. Representation of the frequency, intensity, tracks, genesis, and lysis of LPSs

We found that the simulated seasonal frequency of LPSs in all S2S models was smaller than in ERA-I and MERRA-2, with

the NCEP model having the largest frequency and the BoM model having the smallest frequency. While examining the probability density of the intensity of LPSs, we found that all S2S models had a modal 850-hPa relative vorticity in the range $1\text{--}2 \times 10^{-5} \text{ s}^{-1}$. In MERRA-2, the largest probability density was found for intensity in the range $3\text{--}4 \times 10^{-5} \text{ s}^{-1}$, which suggests that there are stronger LPSs in this reanalysis than in ERA-I and S2S models.

We defined strong LPSs (SLPSs) as systems featuring minimum intensity (850-hPa relative vorticity) greater than or equal to the third quartile intensity of all LPSs in an S2S model or reanalysis datasets and examined their track density, genesis and lysis given their role in triggering high-impact flood events in the Indian subcontinent. We found that all 11 S2S models including the multimodel mean represented transits of SLPSs over the head of the Bay of Bengal and adjoining land area; the NCEP, UKMO, CNRM, ECMWF, and KMA models had the best performance, whereas the BoM, JMA, and CMA models exhibited larger biases in their tracks. The observed west-northwest propagation of SLPSs was not simulated by the CMA model since the midtropospheric steering winds were easterly over the head of the Bay of Bengal and central India. Tracks over west-central India were not simulated by the HMCR model due to faster lysis of SLPS and their low count. We also found that most S2S models as well as the multimodel mean correctly simulated the primary genesis region over the head of the Bay of Bengal and adjoining land area as well as the primary lysis region over eastern India. All the 11 S2S models had fewer genesis and lysis points than ERA-I and MERRA-2.

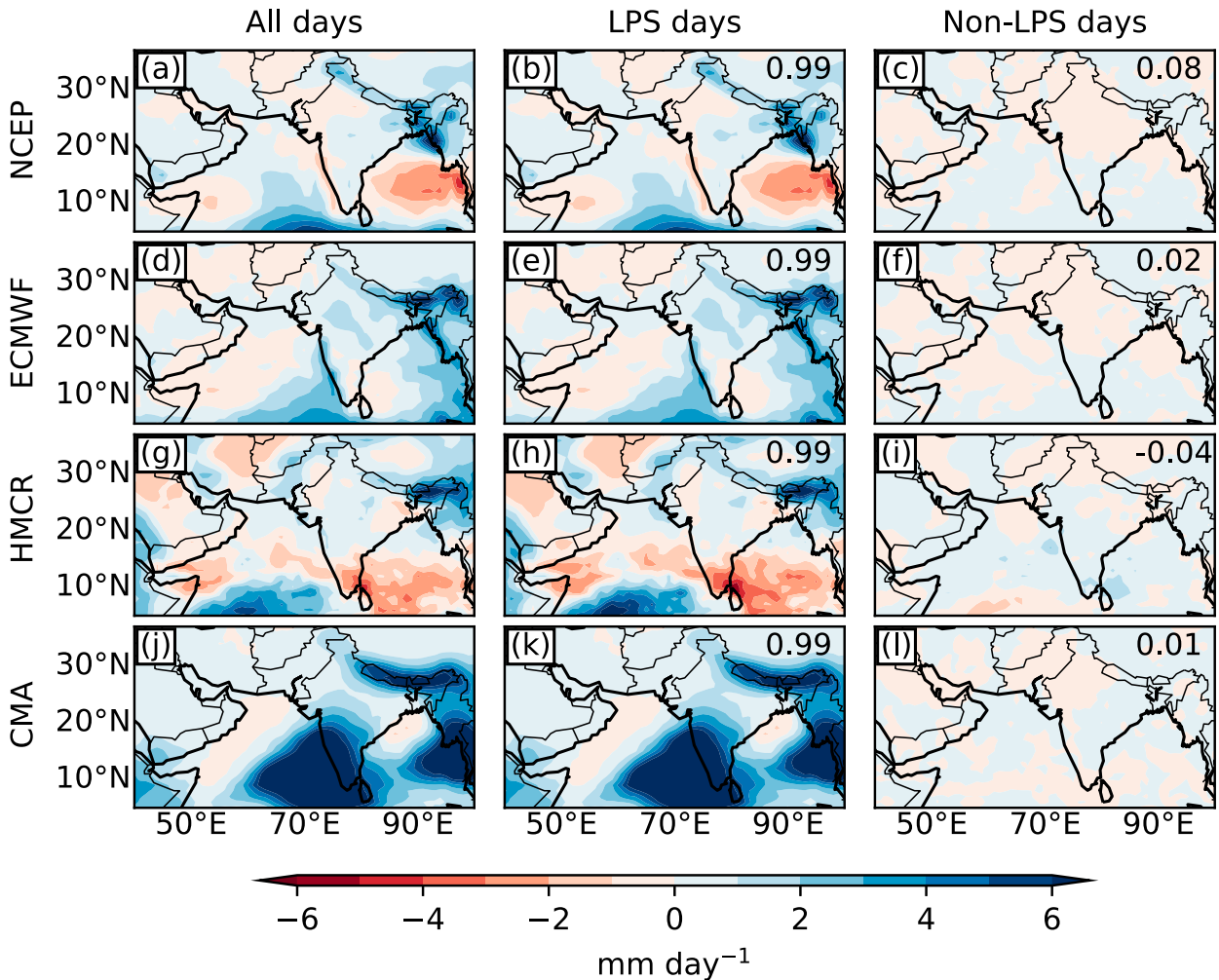


FIG. A3. As in Fig. 11, but for the (a)–(c) NCEP, (d)–(f) ECMWF, (g)–(i) HMCR, and (j)–(l) CMA models.

b. LPS position and intensity error

We investigated the position and intensity errors of LPSs; the CMA model had the largest position error, whereas the ECCO, UKMO, JMA, and KMA models had smaller position errors than most S2S models when LPSs were verified against ERA-I. These models had similar performance when verified against MERRA-2, but a reduction in the position error in the CMA model was seen due to a greater number of westward moving LPSs in MERRA-2 than in ERA-I. We found that the range of the position error was similar in the multimodel mean when LPSs were verified against both reanalysis datasets. Many S2S models including the multimodel mean exhibited a positive bias in the intensity of LPSs at all lead times except the HMCR model when verified against ERA-I. However, most models including the multimodel mean underestimated the intensity when LPSs were verified against MERRA-2. In both cases, the bias was the largest for the BoM and HMCR models and the overall pattern of biases among most models showed consistency. We found that position errors increased with forecast lead time over most of the domain in all S2S models

when verified against ERA-I—the errors were larger over the Arabian Sea than the Bay of Bengal at lead times greater than 12 days. However, this was not seen when LPSs were verified against MERRA-2, suggesting that the presence of more LPSs over the Arabian Sea in MERRA-2 than in ERA-I reduced the position error.

c. The statistical reliability of S2S models at predicting LPSs

A good ensemble spread–error relationship is a desirable property of an ensemble prediction system (EPS) since it indicates the correct representation of all possible forecast outcomes as well as the ensemble mean error by the ensemble spread. We found that irrespective of the reanalysis used for verification, the ECMWF and UKMO models had the best ensemble spread–error relationship for the position and intensity of LPSs, whereas the HMCR model had the worst, possibly due to the perturbation method used in the model. We found that most models were underdispersive for the position and intensity of LPSs—models were more underdispersive for the intensity than the position. These findings are similar to

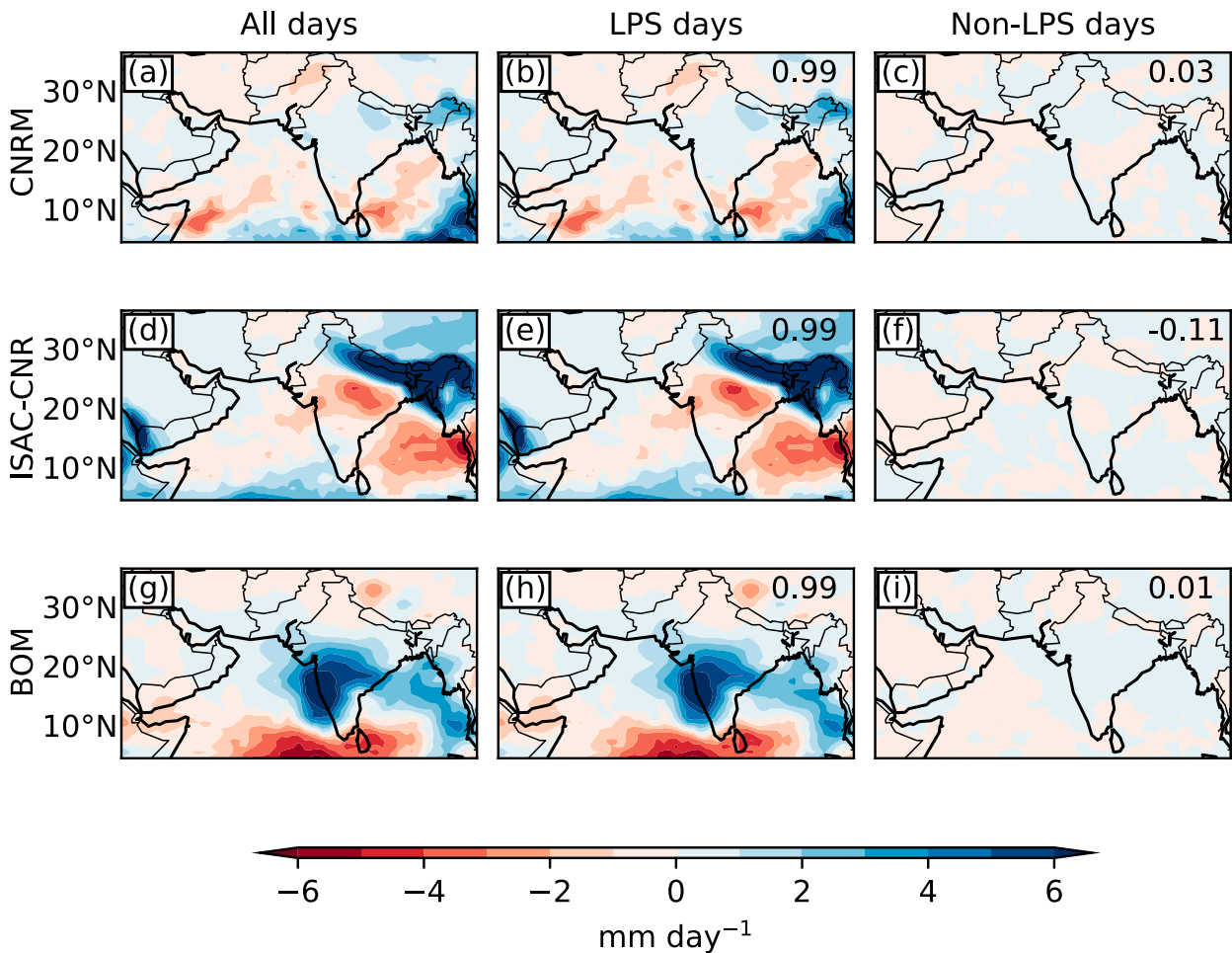


FIG. A4. As in Fig. 11, but for (a)–(c) CNRM, (d)–(f) ISAC-CNR, and (g)–(i) BoM models.

those for extratropical cyclones in the TIGGE dataset (Froude 2010, 2011). Froude (2010, 2011) suspected that the best performance of the TIGGE-ECMWF model at predicting extratropical cyclones was due to a bias toward the ECMWF analysis used for verification in their studies. However, using ERA-I for verification did not give an advantage to the ECMWF model in our study.

d. The influence of forecast lead time and LPSs on the pattern of precipitation errors

In the final phase of this study, we examined the role of forecast lead time and LPSs in influencing precipitation errors in S2S models. The growth of precipitation errors was considered by subtracting GPM IMERG precipitation from S2S precipitation for forecast lead times of 12–15 days minus 0–3 days. We found that S2S models, excluding the KMA, UKMO, and ISAC-CNR models, exhibited a wet bias over most of the monsoon core zone, thereby suggesting an increase in precipitation error with forecast lead time. Models such as CMA exhibited a strong wet bias (up to 20 mm day^{-1}) over the western coast of India, which was related to the intensification of an offshore trough. We also found that the presence of LPSs

influenced the pattern of precipitation errors in all 11 models since there was a strong positive pattern correlation between precipitation errors on all days, and those during the presence of LPSs.

This study opens a new realm of exploring the predictability of LPSs on the time scales of numerical weather prediction models and the extended range and contributes to over a century of literature that has primarily looked at structural and dynamical aspects of LPSs. The results of this paper are potentially useful to meteorologists and disaster management organizations. The most intense LPS related precipitation occurs within $\sim 1000 \text{ km}$ from the LPS center. On several occasions, the presence of LPSs have forced dam operators to suddenly release dam water, thereby triggering dangerous floods such as the 2018 Kerala flood (Lal et al. 2020). Thus, an accurate prediction of an LPS track is crucial to issue flood warnings and skillful forecast of LPSs at longer lead times can help in improving flood forecasts and reservoir operations. Our study presents the first ever evaluation of the prediction of LPSs as well as their precipitation biases, which was a major gap in the literature. Hence, we expect our results to encourage researchers to carry out investigations on improving flood

forecasting in India. Such results will ultimately benefit flood forecasters and dam operators in developing an advanced flood warning system. Further work is required to gain more understanding of factors including the structure of LPSs in the S2S dataset that can influence the predictability of these weather systems. In addition, the contribution of individual models to the multimodel mean results discussed in this study needs to be explored. Multimodel ensembles help in improving the skill of weather forecasts by allowing better estimation of factors such as the forecast uncertainty (Pegion et al. 2019). However, some models can contribute a greater number of older reforecasts to the multimodel ensemble than others due to different initialization dates.

Acknowledgments. AD is funded through the Scholarship for Higher Education in Foreign Countries program, an initiative of the Department of Higher and Technical Education, Government of Maharashtra, India. KMRH is supported through the Weather and Climate Science for Service Partnership (WCSSP) India, a collaborative initiative between the Met Office, supported by the U.K. government's Newton Fund and the Indian Ministry of Earth Sciences (MoES). AGT acknowledges the INCOMPASS project funded by the Natural Environment Research Council (Grant NE/P003117/1). This work is based on the S2S dataset. S2S is a joint initiative of the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP). The original S2S database is hosted at ECMWF as an extension of the TIGGE database.

Data availability statement. The track dataset of LPSs identified in S2S models, ERA-Interim, and MERRA-2 can be accessed at <http://doi.org/10.5281/zenodo.4659797>. All S2S data used in this study are included in Vitart et al. (2017) and are publicly available to research community from the European Centre for Medium-Range Weather Forecasts (ECMWF) at <https://apps.ecmwf.int/datasets/data/s2s-realtime-instantaneous-accum-ecmf/>. The ERA-Interim reanalysis and GPM IMERG datasets are also publicly available at <https://gpm.nasa.gov/data-access/downloads/gpm>.

APPENDIX

Errors in the Position and Intensity of Monsoon Low Pressure Systems and Difference in Daily Precipitation

Figures A1–A2 show errors in the position and intensity of monsoon low pressure systems, respectively, for each S2S model when verified against ERA-Interim. Figures A3–A4 show the difference in daily precipitation for seven S2S models.

REFERENCES

- Boos, W., J. Hurley, and V. Murthy, 2015: Adiabatic westward drift of Indian monsoon depressions. *Quart. J. Roy. Meteor. Soc.*, **141**, 1035–1048, <https://doi.org/10.1002/qj.2454>.
- Buizza, R., P. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, **133**, 1076–1097, <https://doi.org/10.1175/MWR2905.1>.
- Camargo, S. J., 2013: Global and regional aspects of tropical cyclone activity in the CMIP5 models. *J. Climate*, **26**, 9880–9902, <https://doi.org/10.1175/JCLI-D-12-00549.1>.
- , and S. E. Zebiak, 2002: Improving the detection and tracking of tropical cyclones in atmospheric general circulation models. *Wea. Forecasting*, **17**, 1152–1162, [https://doi.org/10.1175/1520-0434\(2002\)017<1152:ITDATO>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<1152:ITDATO>2.0.CO;2).
- Camp, J., M. Roberts, C. MacLachlan, E. Wallace, L. Hermanson, A. Brookshaw, A. Arribas, and A. Scaife, 2015: Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 2206–2219, <https://doi.org/10.1002/qj.2516>.
- Daggupaty, S. M., and D. R. Sikka, 1977: On the vorticity budget and vertical velocity distribution associated with the life cycle of a monsoon depression. *J. Atmos. Sci.*, **34**, 773–792, [https://doi.org/10.1175/1520-0469\(1977\)034<0773:OTVBAV>2.0.CO;2](https://doi.org/10.1175/1520-0469(1977)034<0773:OTVBAV>2.0.CO;2).
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Deoras, A., K. M. R. Hunt, and A. G. Turner, 2021: Track dataset of Indian monsoon low-pressure systems in Subseasonal-to-Seasonal prediction models, ERA-Interim and MERRA-2 reanalysis datasets. Zenodo, <http://doi.org/10.5281/zenodo.4659797>.
- Ding, Y., X. Fu, and B. Zhang, 1984: Study of the structure of a monsoon depression over the Bay of Bengal during summer MONEX. *Adv. Atmos. Sci.*, **1**, 62–75, <https://doi.org/10.1007/BF03187617>.
- Eliot, J., 1884: Account of southwest monsoon storms generated in the Bay of Bengal during 1877–1881. *Memo. Ind. Meteor. Dept.*, **2**, 217–448.
- Francis, P., and S. Gadgil, 2006: Intense rainfall events over the west coast of India. *Meteor. Atmos. Phys.*, **94**, 27–42, <https://doi.org/10.1007/s00703-005-0167-2>.
- Froude, L. S., 2010: TIGGE: Comparison of the prediction of Northern Hemisphere extratropical cyclones by different ensemble prediction systems. *Wea. Forecasting*, **25**, 819–836, <https://doi.org/10.1175/2010WAF2222326.1>.
- , 2011: TIGGE: Comparison of the prediction of Southern Hemisphere extratropical cyclones by different ensemble prediction systems. *Wea. Forecasting*, **26**, 388–398, <https://doi.org/10.1175/2010WAF2222457.1>.
- , L. Bengtsson, and K. I. Hodges, 2007a: The predictability of extratropical storm tracks and the sensitivity of their prediction to the observing system. *Mon. Wea. Rev.*, **135**, 315–333, <https://doi.org/10.1175/MWR3274.1>.
- , —, and —, 2007b: The prediction of extratropical storm tracks by the ECMWF and NCEP ensemble prediction systems. *Mon. Wea. Rev.*, **135**, 2545–2567, <https://doi.org/10.1175/MWR3422.1>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Godbole, R. V., 1977: The composite structure of the monsoon depression. *Tellus*, **29**, 25–40, <https://doi.org/10.3402/tellusa.v29i1.11327>.
- Hodges, K., and R. Emerton, 2015: The prediction of Northern Hemisphere tropical cyclone extended life cycles by the ECMWF ensemble and deterministic prediction systems. Part I: Tropical cyclone stage. *Mon. Wea. Rev.*, **143**, 5091–5114, <https://doi.org/10.1175/MWR-D-13-00385.1>.
- , and N. Klingaman, 2019: Prediction errors of tropical cyclones in the western North Pacific in the Met Office global

- forecast model. *Wea. Forecasting*, **34**, 1189–1209, <https://doi.org/10.1175/WAF-D-19-0005.1>.
- , A. Cobb, and P. L. Vidale, 2017: How well are tropical cyclones represented in reanalysis datasets? *J. Climate*, **30**, 5243–5264, <https://doi.org/10.1175/JCLI-D-16-0557.1>.
- Hopson, T., 2014: Assessing the ensemble spread–error relationship. *Mon. Wea. Rev.*, **142**, 1125–1142, <https://doi.org/10.1175/MWR-D-12-00111.1>.
- Hoskins, B. J., and K. I. Hodges, 2002: New perspectives on the Northern Hemisphere winter storm tracks. *J. Atmos. Sci.*, **59**, 1041–1061, [https://doi.org/10.1175/1520-0469\(2002\)059<1041:NPOTNH>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<1041:NPOTNH>2.0.CO;2).
- Huffman, G. J., D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, and S.-H. Yoo, 2015: NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 4, 26 pp.
- Hunt, K. M. R., and D. J. Parker, 2016: The movement of Indian monsoon depressions by interaction with image vortices near the Himalayan wall. *Quart. J. Roy. Meteor. Soc.*, **142**, 2224–2229, <https://doi.org/10.1002/qj.2812>.
- Hunt, K. M., and J. K. Fletcher, 2019: The relationship between Indian monsoon rainfall and low-pressure systems. *Climate Dyn.*, **53**, 1859–1871, <https://doi.org/10.1007/s00382-019-04744-x>.
- , and A. Menon, 2020: The 2018 Kerala floods: A climate change perspective. *Climate Dyn.*, **54**, 2433–2446, <https://doi.org/10.1007/s00382-020-05123-7>.
- , A. G. Turner, P. M. Inness, D. E. Parker, and R. C. Levine, 2016: On the structure and dynamics of Indian monsoon depressions. *Mon. Wea. Rev.*, **144**, 3391–3416, <https://doi.org/10.1175/MWR-D-15-0138.1>.
- , —, and L. C. Shaffrey, 2018: The evolution, seasonality and impacts of western disturbances. *Quart. J. Roy. Meteor. Soc.*, **144**, 278–290, <https://doi.org/10.1002/qj.3200>.
- Hurley, J. V., and W. R. Boos, 2015: A global climatology of monsoon low-pressure systems. *Quart. J. Roy. Meteor. Soc.*, **141**, 1049–1064, <https://doi.org/10.1002/qj.2447>.
- Joseph, S., and Coauthors, 2015: North Indian heavy rainfall event during June 2013: Diagnostics and extended range prediction. *Climate Dyn.*, **44**, 2049–2065, <https://doi.org/10.1007/s00382-014-2291-5>.
- Kikuchi, K., and B. Wang, 2010: Formation of tropical cyclones in the northern Indian Ocean associated with two types of tropical intraseasonal oscillation modes. *J. Meteor. Soc. Japan*, **88**, 475–496, <https://doi.org/10.2151/jmsj.2010-313>.
- Lal, P., A. Prakash, A. Kumar, P. K. Srivastava, P. Saikia, A. Pandey, P. Srivastava, and M. Khan, 2020: Evaluating the 2018 extreme flood hazard events in Kerala, India. *Remote Sens. Lett.*, **11**, 436–445, <https://doi.org/10.1080/2150704X.2020.1730468>.
- Lee, C.-Y., S. J. Camargo, F. Vitart, A. H. Sobel, and M. K. Tippett, 2018: Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. *Wea. Forecasting*, **33**, 967–988, <https://doi.org/10.1175/WAF-D-17-0165.1>.
- Leith, C., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418, [https://doi.org/10.1175/1520-0493\(1974\)102<0409:TSOMCF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2).
- Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *J. Comput. Phys.*, **227**, 3515–3539, <https://doi.org/10.1016/j.jcp.2007.02.014>.
- Liu, Z., 2016: Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) monthly precipitation products: Initial results. *J. Hydrometeorol.*, **17**, 777–790, <https://doi.org/10.1175/JHM-D-15-0068.1>.
- Murakami, H., 2014: Tropical cyclones in reanalysis data sets. *Geophys. Res. Lett.*, **41**, 2133–2141, <https://doi.org/10.1002/2014GL059519>.
- Pegion, K., and Coauthors, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043–2060, <https://doi.org/10.0.4.151/BAMS-D-18-0270.1>.
- Prakash, S., A. K. Mitra, D. Pai, and A. AghaKouchak, 2016: From TRMM to GPM: How well can heavy rainfall be detected from space? *Adv. Water Resour.*, **88**, 1–7, <https://doi.org/10.1016/j.advwatres.2015.11.008>.
- , —, A. AghaKouchak, Z. Liu, H. Norouzi, and D. Pai, 2018: A preliminary assessment of GPM-based multisatellite precipitation estimates over a monsoon dominated region. *J. Hydrol.*, **556**, 865–876, <https://doi.org/10.1016/j.jhydrol.2016.01.029>.
- Rajeevan, M., S. Gadgil, and J. Bhate, 2010: Active and break spells of the Indian summer monsoon. *J. Earth Syst. Sci.*, **119**, 229–247, <https://doi.org/10.1007/s12040-010-0019-4>.
- Rastogi, D., M. Ashfaq, L. R. Leung, S. Ghosh, A. Saha, K. Hodges, and K. Evans, 2018: Characteristics of Bay of Bengal monsoon depressions in the 21st century. *Geophys. Res. Lett.*, **45**, 6637–6645, <https://doi.org/10.1029/2018GL078756>.
- Ray, K., P. Pandey, C. Pandey, A. Dimri, and K. Kishore, 2019: On the recent floods in India. *Curr. Sci.*, **117**, 204, <https://doi.org/10.18520/cs/v117/i2/204-218>.
- Richardson, D., R. Buizza, and R. Hagedorn, 2005: Final report of the 1st Workshop on the THORPEX Interactive Grand Global Ensemble (TIGGE). WMO TD-1273, WWRP-THORPEX Rep. 5, 39 pp.
- Sandeep, S., R. Ajayamohan, W. R. Boos, T. Sabin, and V. Praveen, 2018: Decline and poleward shift in Indian summer monsoon synoptic activity in a warming climate. *Proc. Natl. Acad. Sci. USA*, **115**, 2681–2686, <https://doi.org/10.1073/pnas.1709031115>.
- Sarkar, R., and A. Choudhary, 1988: A diagnostic study of monsoon depression. *Mausam*, **39**, 9–18.
- Scott, D. W., 2015: *Multivariate Density Estimation: Theory, Practice, And Visualization*. John Wiley & Sons, 384 pp.
- Sørland, S. L., and A. Sorteberg, 2016: Low-pressure systems and extreme precipitation in central India: Sensitivity to temperature changes. *Climate Dyn.*, **47**, 465–480, <https://doi.org/10.1007/s00382-015-2850-4>.
- Sperber, K., H. Annamalai, I.-S. Kang, A. Kitoh, A. Moise, A. Turner, B. Wang, and T. Zhou, 2013: The Asian summer monsoon: An intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. *Climate Dyn.*, **41**, 2711–2744, <https://doi.org/10.1007/s00382-012-1607-6>.
- Tan, J., G. J. Huffman, D. T. Bolvin, and E. J. Nelkin, 2019: IMERG V06: Changes to the morphing algorithm. *J. Atmos. Oceanic Technol.*, **36**, 2471–2482, <https://doi.org/10.1175/JTECH-D-19-0114.1>.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, [https://doi.org/10.1175/1520-0477\(1993\)074<2317:EFANTG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2).
- , and —, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319, [https://doi.org/10.1175/1520-0493\(1997\)125<3297:EFANAT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2).
- Vishnu, S., W. Boos, P. Ullrich, and T. O'Brien, 2020: Assessing historical variability of South Asian monsoon lows and depressions with an optimized tracking algorithm. *J. Geophys. Res. Atmos.*, **125**, e2020JD032977, <https://doi.org/10.1029/2020JD032977>.

- Vitart, F., 2017: Madden–Julian Oscillation prediction and teleconnections in the S2S database. *Quart. J. Roy. Meteor. Soc.*, **143**, 2210–2220, <https://doi.org/10.1002/qj.3079>.
- , and Coauthors, 2017: The subseasonal to seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, <https://doi.org/10.1175/BAMS-D-16-0017.1>.
- Wang, C., G. Tang, Z. Han, X. Guo, and Y. Hong, 2018: Global intercomparison and regional evaluation of GPM IMERG Version-03, Version-04 and its latest Version-05 precipitation products: Similarity, difference and improvements. *J. Hydrol.*, **564**, 342–356, <https://doi.org/10.1016/j.jhydrol.2018.06.064>.
- Yianilos, P. N., 1993: Data structures and algorithms for nearest neighbor search in general metric spaces. *Soda*, **93**, 311–321.
- Yoon, J.-H., and T.-C. Chen, 2005: Water vapor budget of the Indian monsoon depression. *Tellus*, **57A**, 770–782, <https://doi.org/10.3402/tellusa.v57i5.14737>.